TESIS DE LA UNIVERSIDAD
DE ZARAGOZA

2014

3

Eva María Cavero Racaj

Design and evaluation of echocardiograms codification and transmission for Teleradiology systems

Departamento

Instituto de Investigación en Ingeniería [I3A]

Director/es

Alesanco Iglesias, Álvaro
García Moros, José

ISSN 2254-7606

Prensas de la Universidad
Universidad Zaragoza

Tesis Doctoral

DESIGN AND EVALUATION OF ECHOCARDIOGRAMS CODIFICATION AND TRANSMISSION FOR TELERADIOLOGY SYSTEMS

Autor

Eva María Cavero Racaj

Director/es

Alesanco Iglesias, Álvaro
García Moros, José

UNIVERSIDAD DE ZARAGOZA
Instituto de Investigación en Ingeniería [I3A]

2013

DESIGN AND EVALUATION OF ECHOCARDIOGRAMS CODIFICATION AND TRANSMISSION FOR TELECARDIOLOGY SYSTEMS

Eva Mª Cavero Racaj

Supervisors:

Álvaro Alesanco Iglesias and José García Moros

Aragón Institute of Engineering Research (I3A)
Communications Technology Group (GTC)

PhD Dissertation

Biomedical Engineering Doctoral Program

Zaragoza, *October 2013*

A mis padres, Amalia y Santiago

Resumen y conclusiones

Las enfermedades cardiovasculares son la mayor causa de muerte en el mundo. Aunque la mayoría de muertes por cardiopatías se puede evitar, si las medidas preventivas no son las adecuadas el paciente puede fallecer. Es por esto, que el seguimiento y diagnóstico de pacientes con cardiopatías es muy importante. Numerosos son las pruebas médicas para el diagnostico y seguimiento de enfermedades cardiovasculares, siendo los ecocardiogramas una de las técnicas más ampliamente utilizada. Un ecocardiograma consiste en la adquisición de imágenes del corazón mediante ultrasonidos. Presenta varias ventajas con respecto otras pruebas de imagen: no es invasiva, no produce radiación ionizante y es barata. Por otra parte, los sistemas de telemedicina han crecido rápidamente ya que ofrecen beneficios de acceso a los servicios médicos, una reducción del coste y una mejora de la calidad de los servicios. La telemedicina proporciona servicios médicos a distancia. Estos servicios son de especial ayuda en casos de emergencia médica y para áreas aisladas donde los hospitales y centros de salud están alejados. Los sistemas de tele-cardiología pueden ser clasificados de acuerdo al tipo de pruebas. En esta Tesis nos hemos centrado en los sistemas de tele-ecocardiografia, ya que los ecocardiogramas son ampliamente usados y presentan el mayor reto al ser la prueba médica con mayor flujo de datos.

Los mayores retos en los sistemas de tele-ecocardiografia son la compresión y la transmisión garantizando que el mismo diagnóstico es posible tanto en el ecocardiograma original como en el reproducido tras la compresión y transmisión. Los ecocardiogramas deben ser comprimidos tanto para su almacenamiento como para su transmisión ya que estos presentan un enorme flujo de datos que desbordaría el espacio de almacenamiento y no se podría transmitir eficientemente por las redes actuales. Sin embargo, la compresión produce pérdidas que pueden llevar a un diagnostico erróneo de los ecocardiogramas comprimidos. En el caso de que las pruebas ecocardiograficas quieran ser guardadas, una compresión clínica puede ser aplicada previa al almacenamiento. Esta compresión clínica consiste en guardar las partes del ecocardiograma que son importantes para el diagnóstico, es decir, ciertas imágenes y pequeños vídeos del corazón en movimiento que contienen de 1 a 3 ciclos cardiacos. Esta compresión clínica no puede ser aplicada en el caso de transmisión en tiempo real, ya que es el cardiólogo especialista quien debe realizar la compresión clínica y éste se encuentra en recepción, visualizando el echocardiograma transmitido. En cuanto a la transmisión, las redes sin cables presentan un mayor reto que las redes cableadas. Las redes sin cables tienen un ancho de banda limitado, son propensas a errores y son variantes en tiempo lo que puede resultar problemático cuando el ecocardiograma quiere ser transmitido en tiempo real. Además, las redes sin cables han experimentado un gran desarrollo gracias a que permiten un mejor acceso y movilidad, por lo que pueden ofrecer un mayor servicio que las redes cableadas.

Dos tipos de sistemas se pueden distinguir acorde a los retos que presenta cada uno de ellos: los sistemas de almacenamiento y reenvió y los sistemas de tiempo real.

- Los sistemas de almacenamiento y reenvió consisten en la adquisición, almacenamiento y el posterior envió del ecocardiograma sin requerimientos temporales. Una compresión clínica puede ser llevada a cabo previa al almacenamiento. Además de la compresión clínica, una compresión con pérdidas es recomendada para reducir el espacio de almacenamiento y el tiempo de envío, pero sin perder l ainformación diagnóstica de la prueba. En cuanto a la transmisión, al no haber requerimientos temporales, la transmisión no presenta ninguna dificultad. Cualquier protocolo de transmisión fiable puede ser usado para no perder calidad en la imagen debido a la transmisión. Por lo tanto, para estos sistemas sólo nos hemos centrado en la codificación de los ecocardiogramas.

- Los sistemas de tiempo real consisten en la transmisión del ecocardiograma al mismo tiempo que éste es adquirido. Dado que el envío de video clínico es una de las aplicaciones que más demanda en ancho de banda y retardo, la compresión para la transmisión es requerida, pero manteniendo la calidad diagnóstica de la imagen. La transmisión en canales sin cables puede ser afectada por errores que distorsionan la calidad del ecocardiograma reconstruido en recepción. Por lo tanto, métodos de control de errores son requeridos para minimizar los errores de transmisión y el retardo introducido. Sin embargo, aunque el ecocardiograma sea visualizado con errores debido a la transmisión, esto no implica que el diagnóstico no sea posible.

Dados los retos previamente descritos, las siguientes soluciones para la evaluación clínica, compresión y transmisión han sido propuestas:

- Para garantizar que el ecocardiograma es visualizado sin perder información diagnóstica 2 tests han sido diseñados. El primer test define recomendaciones para la compresión de los ecocardiogramas. Consiste en dos fases para un ahorro en el tiempo de realización, pero sin perder por ello exactitud en el proceso de evaluación. Gracias a este test el ecocardiograma puede ser comprimido al máximo sin perder calidad diagnóstica y utilizando así más eficientemente los recursos. El segundo test define recomendaciones para la visualización del ecocardiograma. Este test define rangos de tiempo en los que el ecocardiograma puede ser visualizado con inferior calidad a la establecida en el primer test. Gracias a este test se puede saber si el ecocardiograma es visualizado sin pérdida de calidad diagnóstica cuando se introducen errores en la visualización, sin la necesidad de realizar una evaluación para cada video transmitido o diferentes condiciones de canal. Además, esta metodología para la evaluación clínica de imágenes médicas, puede ser aplicada para otras técnicas de diagnóstico por imagen.

- Para la compresión de ecocardiogramas dos métodos de compresión han sido diseñados, uno para el almacenamiento y otro para la transmisión. Diferentes propuestas son diseñadas, ya que los ecocardiogramas para los dos propósitos tienen características diferentes. Para ambos propósitos un método de compresión en la que las facilidades que incorporan los dispositivos de segmentar la imagen y en la que las características de visualización de los ecocardiogramas han sido tenidas en cuenta ha sido diseñado. Para la compresión del ecocardiograma con el propósito de almacenarlo un formato de almacenamiento fácilmente integrable con DICOM basado en regiones y en el que el tipo de datos y la importancia clínica de cada región es tenido en cuenta ha sido diseñado. DICOM es el formato para el almacenamiento y transmisión de imágenes más ampliamente utilizado actualmente. El formato de compresión propuesto supone un ahorra de hasta el 75 % del espacio de almacenamiento con respecto a la compresión con JPEG 2000, actualmente soportado por DICOM, sin perder calidad diagnostica de la imagen. Los ratios de compresión para el formato propuesto dependen de la distribución de la imagen, pero para una base de datos de 105 ecocardiogramas correspondientes a 4 ecógrafos los ratios obtenidos están comprendidos entre 19 y 41. Para la compresión del ecocardiograma con el propósito de la transmisión en tiempo real un método de compresión basado en regiones en el que el tipo de dato y el modo de visualización han sido tenidos en cuenta se ha diseñado. Dos modos de visualización son distinguidos para la compresión de la región con mayor importancia clínica (ultrasonido), los modos de barrido y los modos 2-D. La evaluación clínica diseñada para las recomendaciones de compresión fue llevada a cabo por 3 cardiologos, 9 ecocardiogramas correspondientes a diferentes pacientes y 3 diferentes ecógrafos. Los ratios de transmisión recomendados fueron de 200 kbps para los modos 2-D y de 40 kbps para los modos de barrido. Si se comparan estos resultados con previas soluciones en la literatura un ahorro mínimo de entre 5 % y el 78 % es obtenido dependiendo del modo.

- Para la transmisión en tiempo real del ecocardiograma un protocolo extremo a extremo basada en el método de compresión por regiones ha sido diseñado. Este protocolo llamado ETP de las siglas en inglés Echocardiogram Transmssion Protocol está diseñado para la compresión y transmisión de las regiones independientemente, pudiendo así ofrecer diferentes ratios de compresión y protección de errores para las diferentes regiones de acuerdo a su importancia diagnostica. Por lo tanto, con ETP el ratio de transmisión mínimo recomendado para el método de compresión propuesto puede ser utilizado, usando así eficientemente el ancho de banda y siendo menos sensible a los errores introducidos por la red. ETP puede ser usado en cualquier red, sin embargo, en el caso de que la red introduzca errores se ha diseñado un método de corrección de errores llamado SECM, de las siglas en inglés State Error Control Method. SECM se adapta a las condiciones de canal usando más protección cuando las condiciones empeoran y usando así más eficientemente el ancho de banda. Además, la evaluación clínica diseñada para las recomendaciones de visualización ha sido llevada a cabo con la base de datos de la evaluación previa. De esta forma se puede saber si el ecocardiograma es visualizado sin pérdida diagnostica aunque se produzcan pérdidas de calidad debido a los errores de transmisión.

En esta tesis, por lo tanto, se ha ofrecido una solución para la transmisión en tiempo real y el almacenamiento de ecocardiogramas preservando la información diagnóstica y usando eficientemente los recursos (disco de almacenamiento y ratio de transmisión). Especial soporte se da para la transmisión en redes sin cables, dando soluciones a las limitaciones que estas introducen. Además, las soluciones propuestas han sido probadas y comparadas con otras técnicas con una red de acceso móvil WiMAX, demostrando que el ancho de banda es eficientemente utilizado y que el ecocardiograma es correctamente visualizado de acuerdo con las recomendaciones de visualización dadas por la evaluación clínica.

Abstract

Cardiovascular diseases (CVDs) are the leading cause of death and disability in the world. For this reason, it is very important to have techniques and services for an accurate diagnosis and follow-up of patients with cardiopathies. One of the most widely medical tests to diagnose CVDs is the echocardiogram. An echocardiogram consists of the continuous acquisition of ultrasound images of the heart. However, in some cases access to medical services is not available, for example in isolated and remote regions where there are no medical specialists. In other cases, the response to emergencies is not fast enough. These problems can be solved with telemedicine systems that provide clinical health care at a distance, helping to eliminate distance barriers and improving access to medical services. Nevertheless, the implantation of tele-echocardiography systems is not straightforward. Digital echocardiographic devices produce high data flows that may exceed the channel capability and the storage space. An efficient compression method is extremely important, especially for real-time transmission in wireless networks owing to the fact that they are band-limited, time-varying and error-prone. Improving the compression method allows cost savings since less bandwidth is required and facilitates the arrival of the echocardiogram at the transmitter without loss of diagnostic information. It is very important to have an image with sufficient quality for making an adequate diagnosis. However, the compression and transmission processes modify the original image. Error control methods must be applied when echocardiograms are transmitted over error prone channels in order to reduce the distortion introduced in the visualized image. Furthermore, a clinical evaluation that gives compression and display recommendations is necessary in order to know if the echocardiogram is visualized with guaranteed clinical quality. Consequently, three main objectives have been addressed in order to provide support to tele-echocardiography systems making an efficient use of resources: the design of an evaluation methodology, the design, evaluation and recommendations for use of compression methods for storage and real-time transmission of echocardiograms, and the design of protocols for transmission of echocardiograms in real-time and recommendations for echocardiogram visualization.

A two phase evaluation methodology for the compression and transmission of clinical images has been designed. This methodology is accurate but less burdensome than other tests proposed in the literature. The evaluation consists of two tests. The first provides compression recommendations and the second provides display recommendations. These tests enable the transmission of clinical videos with a minimal transmission rate and establish whether the clinical image is visualized without losing diagnostic information and without the need to carry out other evaluations for each transmitted video.

An image compression format for storage purpose and an echocardiogram compression method

for real-time transmission purpose have been designed taking into account the echocardiogram characteristics and taking advantage of the segmentation facilities of the device to enhance the compression performance without adding complexity to the device. Both methods compress the regions according to its data type, diagnostic information and visualization characteristics. The proposed storage format saves storage space with respect to the conventional image formats incorporated in the DICOM standard. The compression results depend on the acquisition devices, how the image is displayed, and the compression quality. However, the compression ratios obtained for the proposed format ranged from 19 to 41 without losing diagnostic information. The compression method for real-time transmission that takes advantage of the visualization characteristics of the devices has demonstrated that it compresses echocardiograms more efficiently than conventional video codecs. The recommended transmission rates for the ultrasound regions are the following: 200 Kbps for the 2-D and the color Doppler modes, and 40 Kbps for the M and the pulsed/continuous Doppler modes. This saving in the transmission rate leads to better transmission performance, reducing the transmission time and errors introduced by the channel. These results make possible the transmission of echocardiogram videos over 3G wireless networks and beyond.

An Echocardiogram Transmission Protocol (ETP) for end-to-end real-time transmission of echocardiograms compressed by regions and a States Error Control Method (SECM) that adapts to the channel conditions have been designed. ETP transmits the echocardiogram regions separately, and consequently different transmission rates and error control methods can be used depending on the clinical importance of the region and on the network. Therefore, ETP can be used for transmission in any network. The simulated transmissions have demonstrated that by transmitting the echocardiogram by regions, less bandwidth is used and the echocardiogram is visualized with clinical quality for a greater percentage of time when errors occur than without considering the regions and modes. The saving of bandwidth depends on the mode distribution of the echocardiograms. The percentage of time with adequate clinical quality increases between 5 % and 19 % for a mobile network with WiMAX access, representative settings, nine available echocardiograms and without an error control method. SECM uses different error control methods depending on the channel errors. Furthermore, different configurations can be set depending on the data and the network characteristics. SECM has been demonstrated to adapt the bandwidth used to the channel conditions and to guarantee quality on reception. Furthermore, the echocardiograms are visualized with adequate clinical quality for a mobile network with WiMAX access, representative settings and nine available echocardiograms.

The proposed overall system for real-time echocardiogram transmission over wireless networks achieves a saving in the transmitted bandwidth while ensuring all the available echocardiograms are received with adequate clinical quality. This bandwidth saving depends on the mode distributions of the echocardiogram and on the error distribution. For a mobile network with WiMAX access and representative settings, and the available echocardiograms, the transmitted bandwidth is between 154 kbps and 244 kbps. If the results are compared with previous works where clinical videos with similar resolution are transmitted over WiMAX networks, the saving in transmitted bandwidth is higher than 1 Mbps. This bandwidth saving leads to savings in energy, money and transmission time.

Scientific Contributions

Next peer-reviewed scientific publications have been derived from the research of this thesis.

Publications in International Journals (JCR indexed)

- **E. Cavero**, A. Alesanco, and J. Garcia, "Enhanced Protocol for Real-Time Transmission of Echocardiograms Over Wireless Channels," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 11, pp. 3212–3220, Nov. 2012.
- **E. Cavero**, A. Alesanco, L. Castro, J. Montoya, I. Lacambra, and J. Garcia, "SPIHT-Based Echocardiogram Compression: Clinical Evaluation and Recommendations of Use," *IEEE Journal of Biomedical and Health Informatics*, vol. 17, no. 1, pp. 103–112, Jan. 2013.

Submitted Publications in International Journals (JCR indexed)

- **E. Cavero**, A. Alesanco, and J. Garcia, "Real-time Echocardiogram Transmission Protocol based on Regions and Visualization Modes ," *IEEE Journal of Biomedical and Health Informatics.*
- **E. Cavero**, A. Alesanco, and J. Garcia, "Low-complexity Ultrasound Image Format Based on Regions," *IEEE Journal of Biomedical and Health Informatics.*

Preparation Publications in International Journals (JCR indexed)

- **E. Cavero**, A. Alesanco, and J. Garcia, "Real-Time Transmission of Echocardiogram: Display Recommendations and Enhanced Methods for the Transmission based on Visualization Modes,".

Publications in International Conferences and Proceedings

- **E. Cavero**, A. Alesanco, and J. Garcia, "A new approach for echocardiogram compression based on display modes," in *2010 10th IEEE International Conference on Information Technology and Applications in Biomedicine (ITAB)*, Nov. 2010.
- **E. Cavero**, A. Alesanco, and J. Garcia, "Real-time transmission of 2d echocardiograms over wimax networks," in *Computing in Cardiology (CinC), 2012*, Sept. 2012, pp. 997–1000.

- J. Garcia, A. Alesanco, and **E. Cavero**, “Cardiac signals coding and transmission in real-time mobile telecardiology applications,” in *Computing in Cardiology (CinC), 2012*, Sept., pp. 317–320.
- **E. Cavero**, A. Alesanco, L. Trajkovic, C. Pattichis, and J. Garcia, “Proposal of real-time echocardiogram transmission based on visualization modes with wimax access,” in *Computing in Cardiology (CinC), 2013*, Sept.

Publications in National Conferences and Proceedings

- **E. Cavero**, A. Alesanco, and J. Garcia, “Nueva propuesta de compresión para ecocardiogramas basada en modos de visualización (A new proposal for echocardiogram compression based on visualization modes),” in *CASEIB 2010 - XXVIII Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, Nov. 2010.
- **E. Cavero**, A. Alesanco, and J. Garcia, “Nueva propuesta de compresión para el almacenamiento de pruebas ecocardiográficas (A new compression proposal for the storage of echocardiography exams),” in *CASEIB 2011 - XXIX Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, Nov. 2011.
- **E. Cavero**, A. Alesanco, and J. Garcia, “Transmisión en tiempo real de Ecocardiogramas 2D sobre redes WiMAX (Real-time transmission of 2-D echocardiograms over WiMAX),” in *CASEIB 2012 - XXX Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, Nov. 2012.

Contents

List of Figures

List of Tables

Acronyms

ACK Acknowledge

ACR American College of Radiology

ADSL Asymmetric Digital Subscriber Line

ANOVA Analysis of variance

ARQ Automatic Repeat ReQuest

ASCII American Standard Code for Information Interchange

ASE American Society of Echocardiography

AVC Advanced Video Coding

BE Best-Effort

BS Base Station

BWA Broadband Wireless Access

CBR Constant Bite-Rate

CEZW Color Embedded Zero-tree Wavelet

CLD Cross-Layer Design

CSPIHT Color Set Partitioning in Hierarchical Trees

CT Computed Tomography

CVDs Cardiovascular diseases

DICOM Digital Imaging and Communications in Medicine

DS3 Digital Signal 3

DTD Document Type Definition

EAE European Association of Echocardiography

ECG Electrocardiogram

ETP Echocardiogram Transmission Protocol

EZW Embedded Zero-tree Wavelet

FDD Frequency Division Duplex

FEC Forward Error Correction

FMO Flexible Macroblock Ordering

G	Generation
HEVC	High Efficiency Video Coding
HL7	Health Level Seven
HSUPA	High-Speed Uplink Packet Access
ICT	Information and Communication technologies
IEEE	Institute of Electrical and Electronics Engineers
IMT-Anvanced	International Mobile Telecommunications-Advanced
IP	Internet Protocol
ITU-T	International Telegraph Union Telecommunication Standardization Sector
JPEG	Joint Photographic Experts Group
JPIP	JPEG 2000 Interactive Protocol
JVT	Joint Video Team
KLT	KahunenLoève Transform
LIP	List of isolated insignificant pixels
LIS	List of insignificant sets
LSP	List of isolated significant pixels
NEMA	National Electrical Manufacturers Association
nrtPS	Non-real-time Polling Service
MAC	Medium Access Control
MPEG	Moving Picture Experts Group
MSE	Mean Squared Error
OCR	Optical Character Recognition
OFDM	Orthogonal frequency-division multiplexing
PACS	Picture Archiving and Communications Systems
PHY	Physical
PS	Polling Service
PSNR	Peak Signal-to Noise Ration
QoS	Quality of Service
RCVTP	Reliable Clinical Video Transmission Protocol
RLE	Run Length Encoding
ROHC	RObust Header Compression
ROI	Region of Interest
RS	Reed−Solomon
RTP	Real-Time Transport Protocol

rtPS	Real-time Polling Service
RTO	Retransmission Time-Out
RTT	Round-Trip Time
SCV	Scalable Video Coding
SECM	States Error Control Method
SNR	Signal-to-noise ratio
SOFDMA	Scalable Orthogonal Frequency Division Multiple Access
SPIHT	Set Partitioning in Hierarchical Trees
SSs	Subscriber Stations
TDD	Time-Division Duplex
TLS	Transport Layer Security
TCP	Transmission Control Protocol
UDP	User Datagram Protocol
UGS	Unsolicited Grant Service
UTF-8	8-bit Unicode Transformation Format
VCEG	Video Coding Experts Group
WAN	Wide Area Network
WiMAX	Worldwide Interoperability For Microwave Access
WMAN	Wireless Metropolitan Area Networks
XML	eXtensible Markup Language

Chapter 1

Introduction

Cardiovascular diseases (CVDs) are the leading cause of death and disability in the world. For this reason, it is very important to have techniques and services for an accurate diagnosis and follow-up of patients with cardiopathies. One of the most widely medical tests to diagnose CVD is the echocardiogram. In some cases access to medical services is not available or the response to emergencies is not fast enough. These problems can be solved with telemedicine systems that provide clinical health care at a distance, helping to eliminate distance barriers and improving access to medical services. Nevertheless, the implantation of tele-echocardiography systems is not straightforward. Numerous challenges should be addressed in order to completely exploit the advantages that these systems provide. This Chapter is organized as follow. The motivations, and general concepts and definitions such as cardiovascular diseases, cardiac test, telemedicine, eHealth, tele-echoacardiography systems are described in Section 1.1. Section 1.2 addresses the main aspects of the tele-echocardiography systems: echocardiography, challenges and working scenarios. Section 1.3 defines the approach and objectives of this Thesis. Section 1.4 describes the research context. Finally, the Thesis outline is presented in Section 1.5.

1.1 Motivation

CVDs are caused by disorders of the heart and blood vessels, and includes coronary heart disease (heart attacks), cerebrovascular disease (stroke), high blood pressure (hypertension), peripheral artery disease, rheumatic heart disease, congenital heart disease and heart failure. The major causes of cardiovascular disease are tobacco use, physical inactivity, an unhealthy diet and harmful use of alcohol. As a matter of fact, CVDs are the leading causes of death and disability in the world. According to the World Health Organization [1], an estimated 17.3 million people died from CVDs in 2008, representing 30% of all global deaths. Of these deaths, an estimated 7.3 million were due to coronary heart disease and 6.2 million were due to strokes. Moreover, according to the latest report from the American Heart Association [2], in 2009, one in nine death certificates (274601 deaths) in the United States mentioned heart failure and more than 2150 Americans die of CVDs each day, an average of 1 death every 40 seconds.

In Europe, 35.6% of all deaths were due to CVDs in 2010, and 28.3 % in Aragón constituting the leading cause of death according to National Institute of Stadistics [3]. Although a large proportion of CVDs is preventable, fatalities continue to rise mainly because preventive measures are inadequate. For this reason, it is very important to have techniques and services for an accurate diagnosis and follow-up of patients with cardiopathies.

Nowadays, as technology advances, doctors have access to a wider range of medical tests which can be used to diagnose CVDs in a safe and efficient manner. In cardiology, cardiac tests help to confirm a clinical diagnosis as well as assisting in the follow-up of patients with cardiopathies. The main medical tests for CVDs, apart from the physical examination, are the electrocardiogram (ECG), echocardiogram, coronary angiography, computed tomography (CT), exercise stress test, and holter monitoring. One of the most widely used techniques is echocardiography. An echocardiogram is based on the continuous acquisition of ultrasound images of the heart. It has several advantages compared with other medical imaging techniques: it is non invasive, it does not produce ionized radiation and it is cheap.

Several definitions can be found for the term eHealth [4]. Broadly speaking, eHealth can be defined as the delivery of healthcare and the exchange of healthcare information across distances [5] by means of information and communications technologies (ICT) [6]. eHealth is an umbrella term that includes telemedicine, electronic medical records, and other components of health information technology. Telemedicine refers specifically to the provision of clinical services while the term ehealth can refer to clinical and non-clinical services involving medical education, administration, and research.

Telemedicine systems have grown rapidly because they offer several benefits such as improved access to medical services, a reduction in the cost of healthcare and improvements in the quality of services. In [7], an overall review of the effectiveness of telemedicine was presented showing that twenty-one studies concluded that telemedicine works well and has positive effects. These include improved therapeutic effects, increased efficiencies in the services, and greater technical usability. Nowadays, telemedicine systems are commonly used to provide remote assistance, such as remote diagnosis (telediagnosis) and second opinions from specialized physicians (teleconsultation), to medical institutions that do not have specialized human resources. They are also used for the remote monitoring of patients (telemonitoring). These systems are especially helpful in countries where there is no easy access to medical care due to geographical barriers or a widely dispersed distribution of physicians.

Several classifications of telemedicine systems can be made according to the nature of the data (images, video, audio, signal, text), the transmission network (wired and wireless) and the required time (store-and-forward and real-time). Different challenges and problems arise for each telemedicine system according to these classifications, each one requiring specific research and investigation.

One of the most important areas for the application of telemedicine is telecardiology due to the fact that CVDs are the leading cause of death and disability in the world. In general, it is possible to distinguish three types of telecardiology systems according to the clinical test applied: phonocardiogram, ECG and echocardiogram. The latter test is the most comprehensive since it can include the previous two. It is also the most challenging since it represents the biggest data flow which makes compression and transmission more difficult.

Thanks to the development of efficient compression techniques and wideband communication channels [8], low-cost and clinically accurate tele-echocardiography systems can be implemented even for real-time diagnosis. A review of 50 published works on tele-echocardiography appeared in [9]. The analysis showed a generalized widespread increase in the use of digital tele-echocardiography thanks to the development of information technology. This diffusion has sometimes been accompanied by research studies into diagnostic accuracy and cost-benefit analyses with special emphasis on the economic and social impact. It has been shown in [10] that tele-echocardiography diminishes costs and enhances cardiac disease management.

1.2 Tele-Echocardiography

1.2.1 Echocardiography

The echocardiography is the most widely used imaging modality in clinical cardiology. It is based on the continuous acquisition of ultrasound images of the heart and it is the technique of choice for the diagnosis and follow-up of most heart diseases. The echocardiogram allows doctors to diagnose, evaluate, and monitor [11]: abnormal heart valves, atrial fibrillation, congenital heart disease, damage to the heart muscle in patients who have had heart attacks, heart murmurs, infection in the sac around the heart (pericarditis), infection on or around the heart valves (infectious endocarditis), pulmonary hypertension, the pumping function of the heart for people with heart failure and the source of a blood clot after a stroke.

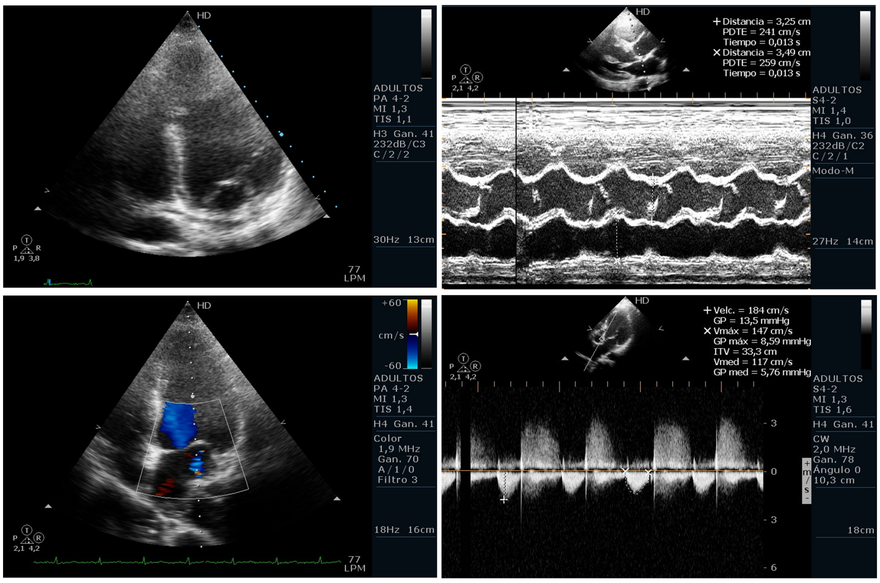


Figure 1.1: *Echocardiogram modes. B mode (top left), color Doppler mode (bottom left), M mode (top right), and pulsed/continuous Doppler mode (bottom right).*

The echocardiograms have different modes of operation. Each mode presents different representation of the heart. There are four basic modes of operation recommended by the European Association of Echocardiography (EAE) [12] and the American Society of Echocardiography (ASE) [13] that are incorporated in basic cardiac ultrasound devices. These are: B, M, color Doppler and pulsed/continuous Doppler (see Figure 1.1). The B mode displays a 2-D image that represents the heart and its movement. The M mode represents a 1-D view of the cardiac structures moving over time. Together, the B and M modes permit us to measure the size, thickness and movement of the heart. The color Doppler mode displays a 2-D image of the heart as the B mode, but shows velocity of the blood-flow by using a color code. The pulsed and continuous mode represents a 1-D view of the blood velocity over time and permits taking velocity measurements in a specific portion of the heart. Recent developments in echocardiography such as color M mode, 3-D echocardiography and second-generation intravenous contrast agents have widened the applications of the technique. However, many of these newer developments do not necessarily have to be currently included in a routine transthoracic echo studies according the EAE recommendations.

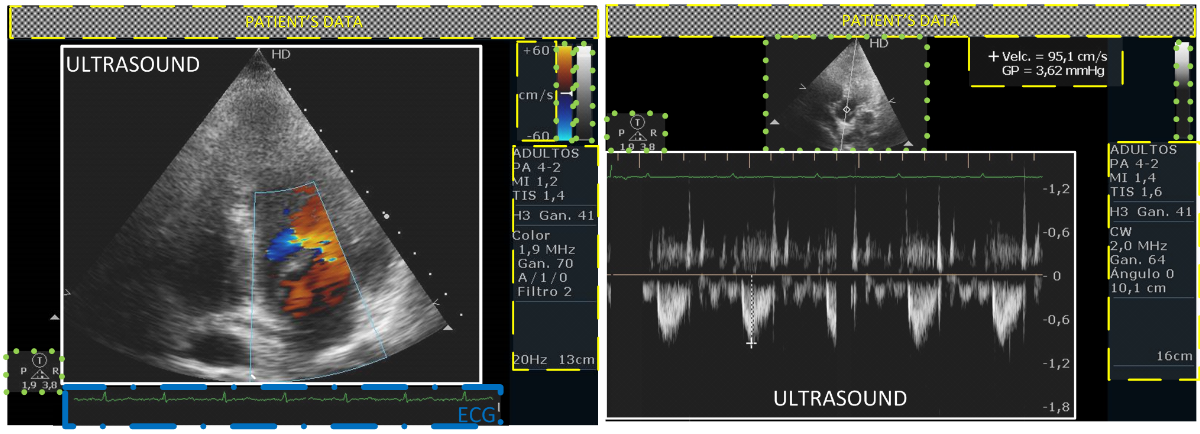


Figure 1.2: *Echocardiogram regions of a color Doppler mode (on the left) and a pulsed Doppler mode (on the right) captured with a Philips Envisor device. The white solid line contains the ultrasound, the blue dotted and dashed line the ECG, the green dotted line the auxiliary images and the yellow dashed line the text.*

The echocardiograms have special characteristics in the way that the image is displayed that should be taken into account in order to design an efficient tele-echocardiography system. As can be seen in Figure 1.2, the echocardiogram has four main types of regions: ultrasound image, ECG, auxiliary images and text. The ultrasound and ECG regions contain the relevant clinical information, they are the regions of interest. Moreover, the echocardiogram may also incorporate sound to listen to the heartbeat.

The accuracy of an echocardiogram depends on the knowledge, skill and experience of both the individual performing the study and the physician interpreting the study. For this reason, according to the ASE [13] a transthoracic echocardiogram must be performed by a physician with Level II training in echocardiography, or the equivalent, or by an accredited sonographer [14]. The echocardiogram should also be interpreted by a physician who has Level II training in echocardiography, or the equivalent. In a typical tele-echocardiography scenario, the person who acquires the echocardiogram, the sonographer, is not able to perform the interpretation and diagnosis of the echocardiogram. This is usually done at another location.

1.2.2 Challenges

According to the EAE recommendations [12] the time allocated for a standard transthoracic study should be at least 30 min. A standard study will therefore require at least 60 GB of uncompressed data, the total size depending on the spatial and temporal resolution of each study. For example, a typical echocardiogram with a frame size of 720 x 576 pixels, 24 bits per pixel and 25 fps presents a raw data flow of about 250 Mbit per second.Furthermore, according to the EAE [12] it is recommended to store the studies for at least the duration of the patients' life expectancy for subsequent analysis and review, although legal requirements affecting the duration of mandatory data storage have to be met. Consequently, compression must be applied for two purposes: reduction of storage requirements and reduction in the transmission rate. Unfortunately, lossless methods can only provide limited compression rates (up to 4:1) [15] which are insufficient. Therefore, lossy compression has to be used since it is able to reduce the data flow considerably, having compression rates up to 20:1 with better quality than the original analog echocardiograms [16] in the sense that noise is reduced. Moreover, in order to compress the echocardiogram efficiently, its visualization characteristics should be taken into account.

Nevertheless, it is very important to note that lossy compression modifies the original video and may decrease its quality. The higher the compression rate, the higher the distortion [17, 18]. In medical images, a minimum quality is required to be able to make an adequate diagnosis. There is a compromise between compression rate and clinical quality. Thus, the minimum recommended compression rates required for the compressed echocardiogram to achieve adequate clinical quality have to be provided. In order to quantify the clinical distortion introduced in the compressed signal, distortion indexes that quantify the real degradation in the diagnosis content of echocardiograms are necessary. A testbed that unifies and reflects the clinical evaluation procedure should be designed.

The integration of the digital echocardiogram replacing traditional video tape recording in echocardiogram machines [19] presented several advantages that were addressed in [20, 21]. The main advantage of the introduction of the digital echocardiogram was the possibility of including clinical compression for storage purpose. Clinical compression stores only the information that is important for the diagnosis. For echocardiograms, clinical compression was defined in [12, 19, 20], and the accuracy of the approximation was demonstrated in [22]. Consequently, instead of storing the whole video, the specialist cardiologist chooses several images of different heart visualizations or several images of the same heart visualization of at least one but preferably three cardiac cycles (from 14 to 64 frames). Clinical compression considerably reduces the storage space, although digital compression is also recommended. It is important to note that clinical compression is not applicable for echocardiogram transmissions in real-time, since it is the specialist cardiologist who makes the clinical compression while watching the transmitted echocardiogram. Thus, the whole echocardiogram has to be transmitted.

Another important aspect to be taken into account is that the echocardiogram image may contain text regions with the patient's personal data (see Figure 1.2). Thus the echocardiogram must be treated with special care since it identifies the patient. Consequently, the protection of these tests during transmission is required by several governmental regulations such as the HIPAA [23] in the U.S., the PIPEDA [24] in Canada, the LOPD [25] in Spain and the Digital Signature Laws in several countries.

Transmission presents more challenges for wireless than for wired channels because the former are band limited, time varying and error prone. This is a particular problem for medical video streaming applications [26]. However, mobile healthcare (m-health) has undergone an impressive development over recent decades [27, 28] thanks to the improvement of wireless technologies. This emerging concept has seen the evolution of e-health systems from traditional desktop telemedicine platforms to wireless and mobile configurations, allowing access to telemedicine services anywhere.

Two types of tele-echocardiogram systems can be distinguished: store&forward and real-time systems. The working procedures and main challenges addressed in this Thesis for each system are described below.

- The **storage&forward systems** involve acquiring, encoding, storing and transmitting the echocardiogram at a convenient time for assessment offline, without time deadline requirements. First, while the echocardiogram is performed, the cardiologist selects the relevant images for the diagnosis (clinical compression). After the clinical compression, the acquisition device provides several images of different heart visualizations and a set of images from the same heart visualization (from 14 to 64 frames) for each echocardiogram instead of a video. Although a clinical compression is performed previous to storage, a lossy compression is also required to reduce the storage space and the transmission time, but without losing diagnosis quality. Since there are no time requirements for transmitting the echocardiograms, transmission is not a difficulty. Any reliable protocol can be used for the transmission of the whole echocardiogram without losing quality. Furthermore, store&forward telemedicine can be used with any bandwidth, although there is a trade-off between transmission time and bandwidth. In conclusion, this Thesis is focused on the encoding aspect of store&forward systems, which presents the most significant challenge.

- The **real-time transmission systems** involve transmitting the echocardiogram video at the same time that it is acquired. Medical video streaming is the most demanding application in terms of bandwidth and delay, and hence requires compression for transmission while at the same time maintaining high image quality on reception in order to avoid losing diagnostic information. A bandwidth saving leads to better transmission performance, reducing the transmission time and error introduced by the channel. In order to provide an accurate diagnosis, it is not only necessary to have a compression method that guarantees clinical quality, but it is also essential to be able to guarantee the integrity of the video during the transmission process. It is well known that wireless channels are error prone. Consequently, digital transmission over wireless channels may be affected by erroneous bits that distort the reconstruction of the video on reception. Hence, error control methods are required in order to minimize the delay and errors in transmission. However, even if the echocardiogram is not visualized all the time with minimal acceptable quality, it may be possible to preserve the diagnostic information. An assessment by expert cardiologist is required in order to find out the percentage of time that the echocardiogram can be visualized with lower than minimum quality but without losing diagnostic quality.

1.2.3 Working Scenarios

With emerging wireless technologies, patients can access healthcare services not only from hospitals, but also from rural healthcare centers, ambulances, ships, trains, airplanes and homes. The tele-echocardiography scenarios dealt with in this Thesis are shown in Figure 1.3 and described below:

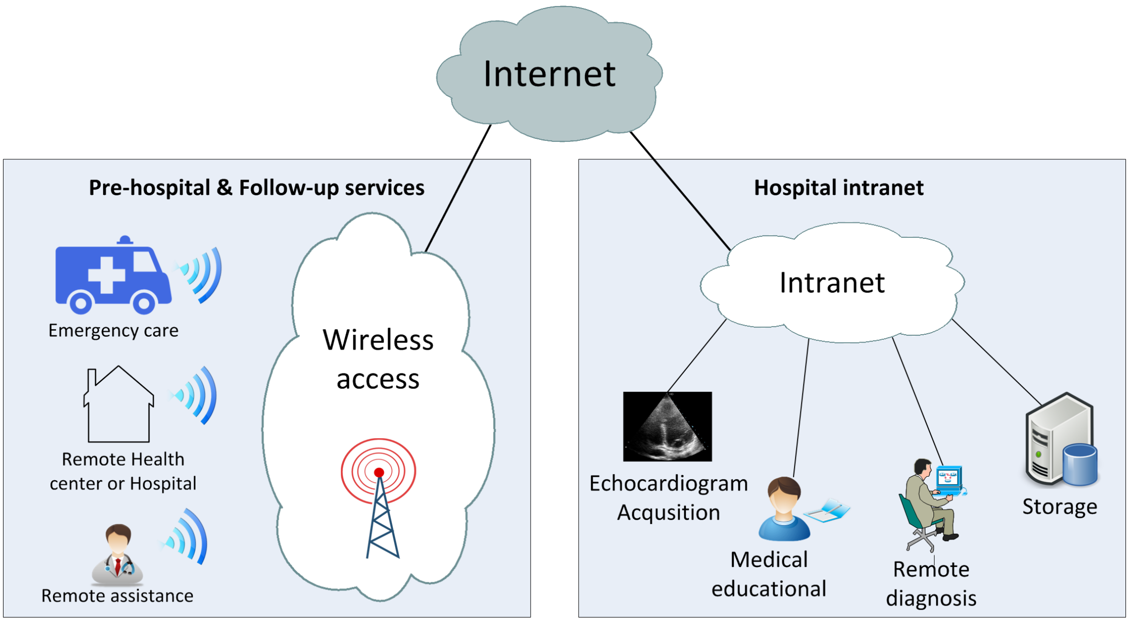


Figure 1.3: *Typical scenarios of tele-echocardiography for real-time and store&forward systems.*

- Telediagnosis of patients in remote areas with wireless access (real-time systems). The patient who lives in a remote area does not have to move to the hospital where the expert cardiologist works for the follow-up and early diagnosis of cardiovascular diseases. The echocardiogram is performed in a remote location by a sonographer and the echocardiogram is sent to the expert cardiologist who visualizes the echocardiogram in real-time and makes the diagnosis.

- Emergency cases (real-time systems). The transmission of the echocardiogram video begins after the patient is transferred into the ambulance or other conveyance. The basic idea is to communicate ultrasound to an emergency physician, to further assist in the diagnosis and to prepare for the patient's admission to the hospital.

- Medical educational and collaborative evaluations (real-time and store&forward systems). This allows centers delivering cardiology education or collaborative lectures to show echocardiogram acquisition in real-time as well as stored echocardiograms to hospitals and health care centers. These applications provide fundamental principles, firsthand knowledge and evidenced-based methods for critical analysis of established clinical practice standards, and comparisons with newer advanced alternatives. Various centers collaborate and share their perspectives based on their location, available staff, and available resources.

- Storage of examinations (store&forward systems). The studies are stored for subsequent analysis and review or for onward transmission to another center in the event that the patient changes location.

1.3 Thesis Approach and Objectives

The general approach of this Thesis is to research and make contributions to the field of ICT applied to the health area. Nowadays, research results in Communication Technologies and the Information Society are considered strategic. Moreover, the application of these technologies to the health area through new telemedicine services facilitates citizens' access to the health system. Therefore, investigations in this area are highly relevant thanks to the potential benefits for patients, doctors and the health system as a whole.

The main aim of this Thesis is to investigate telemedicine systems applied in cardiology environments, since cardiovascular diseases are the leading cause of death in the developing world mainly because they are not diagnosed sufficiently early. Thus, the overall objectives are:

- The design, evaluation and recommendations for use of compression methods for storage and real-time transmission of echocardiograms.
- The design and evaluation of protocols for transmission of echocardiograms in real-time, and recommendations for the echocardiogram visualization.

On a deeper level, further detailed objectives can be mentioned. These are presented as follows, subdivided into the main topics of the Thesis, i.e. compression and transmission. Compression:

- To conduct reviews on the state of the art in general aspects of compression methods for both image and video, and clinical quality. Specifically, to review the literature on compression techniques used in wireless medical video transmission systems.
- To design a compression image format taking into account the characteristics of stored echocardiograms in order to save storage space while preserving clinical quality.
- To design an efficient method to encode the echocardiogram in real-time taking into account its visualization characteristics in order to use less bandwidth while preserving its clinical quality.
- To design an accurate echocardiogram evaluation methodology in order to provide a recommendation for echocardiogram compression.
- To provide recommendations of use for the proposed compression method for real-time transmission and comparisons with the transmission rates described in the literature.

Transmission:

- To conduct reviews on the state of the art in general aspects of wireless technologies and transmission protocols. Specifically, to review the literature on wireless medical video transmission systems.
- To design a protocol to transmit the echocardiogram in real-time taking into account its characteristics in order to accomplish transmission using less bandwidth while preserving clinical quality.
- To design an error control method for echocardiogram transmission that allows adequate diagnosis even over error prone channels.
- To provide display recommendations for an accurate diagnosis of echocardiograms.
- To study the transmission of echocardiograms in real-time over wireless networks with the designed techniques. To study quality parameters such as transmitted bandwidth and diagnostic validation.

1.4 Research Context

This Thesis has been mostly developed within the following funded projects of the Telemedicine research line of the Communications Technologies Group (GTC), belonging to the Aragón Institute of Engineering Research (I3A) of the University of Zaragoza:

1. DGA - PI029/09: "Analysis of echocardiogram coding and real-time transmission through communications networks".
2. MCINN - TIN2008-00933/TSI: "New telemonitoring systems for e-Health services".
3. MCINN - TIN-2011-23792/TSI: "Ontology-based interoperable architecture for patients telemonitoring and clinical decision support".

The research group collaborates with other entities, institutions, universities and hospitals, both national and international. Specially mention can be made of the following international cooperation partners relevant to this Thesis:

- The e-Health Laboratory of the Department of Computer Science, of the University of Cyprus, through Professor Constantinos S. Pattichis, Dean of the School of Pure and Applied Sciences.
- The Communications Network Laboratory in the School of Engineering Science at Simon Fraser University, Burnaby, British Columbia, Canada through Professor Ljiljana Trajkovic.

Two research stages have been done in collaboration with these research groups.

At a national level, specific mention can be made of the following clinical collaborations partners:

- The Cardiology Service of the Lozano Blesa Clinic Hospital (Zaragoza) through Dr. Isaac Lacambra, Lena Castro and José Montoya.
- The Cardiologist Service of the "Hospital viamed Montecanal" (Zaragoza) through Dr. Pedro Serrano.
- The Cardiology Service of the Miguel Servet Hospital (Zaragoza) through Dr. Ernest Spitzer.

1.5 Thesis Outline

The remainder of this Thesis is structured as follows:

- Chapter 2 provides a detailed review of the state of the art of compression of image and video for storage and transmission in real-time purposes, transmission of ultrasound videos, wireless technologies and clinical quality. It also summarizes the most important characteristics of recent ultrasound wireless transmission systems. Finally, improvements are suggested that can be made to existing systems in order to achieve better results in compression, transmission and clinical evaluation.
- Chapter 3 describes echocardiogram databases and characteristics for both storage and real-time purposes. It also contributes a clinical evaluation methodology for compression recommendations of echocardiograms, a compression technique for storage purpose and its results compared with other compression techniques, and finally a compression technique for real-time transmission purposes, including recommendations and compression results. The results are compared with the results of previous systems.
- Chapter 4 describes an evaluation methodology for the display recommendations of medical images after transmission, specifically for echocardiograms that have been encoded with the method proposed in Chapter 3; display recommendations for the echocardiogram after compression for real-time transmission with the proposed method and recommended transmission rates listed in Chapter 3, a protocol for echocardiogram transmission in real-time; an error control method and its configuration for echocardiogram transmission using the proposed protocol, and finally results and discussion for real-time echocardiogram transmission over wireless channels.
- Chapter 5 presents research objectives achieved, contributions and accomplished results of this Thesis, and future lines of research.

Chapter 2

Tele-Echocardiography Realities: Background and What Can Be Improved?

The challenging issues in tele-echocardiography systems dealt with in this thesis are the key factors of encoding algorithms, transmission protocols, error control methods, wireless transmission technologies and clinical quality. The tele-echocardiography system scheme followed in this Thesis is depicted in Figure 2.1. The echocardiogram is acquired and can then be stored or transmitted in real-time. For both purposes, the echocardiogram, in video or image format, has to be compressed with the maximum compression rate while preserving the clinical quality. However, the data acquired for each purpose is different and therefore the compression methods must also be different. In the case of transmitting the echocardiogram, the protocols have to be carefully chosen in order to guarantee reception of the echocardiogram with minimal delay and without losing diagnostic information even if not all the information gets to the receiver in time. Furthermore, the wireless channel has to comply with the system requirements, acceptable delay and transmission rates. Sections 2.1, 2.2, 2.3, 2.4, 2.5 and 2.6 address the above-mentioned key factors in terms of the state of the art, improvements that can be made to achieve better results and the most relevant technologies that have been used in this Thesis. Finally, the most important characteristics of recent ultrasound wireless transmission systems are summarized in Section 2.7 and conclusions relating to the design premises used in this Thesis to improve tele-echocardiography systems are shown in Section 2.8.

2.1 Compression for Storage

Hospitals and medical centers produce large volume of digital medical images that require considerable storage space. Standardization of storage format is critical to enable interoperability within and between centers and equipment from different vendors. A survey of the different medical image formats and compression techniques is presented in [29]. Digital Imaging and Communications in Medicine (DICOM) [30] is the most widely used and accepted standard for effective medical

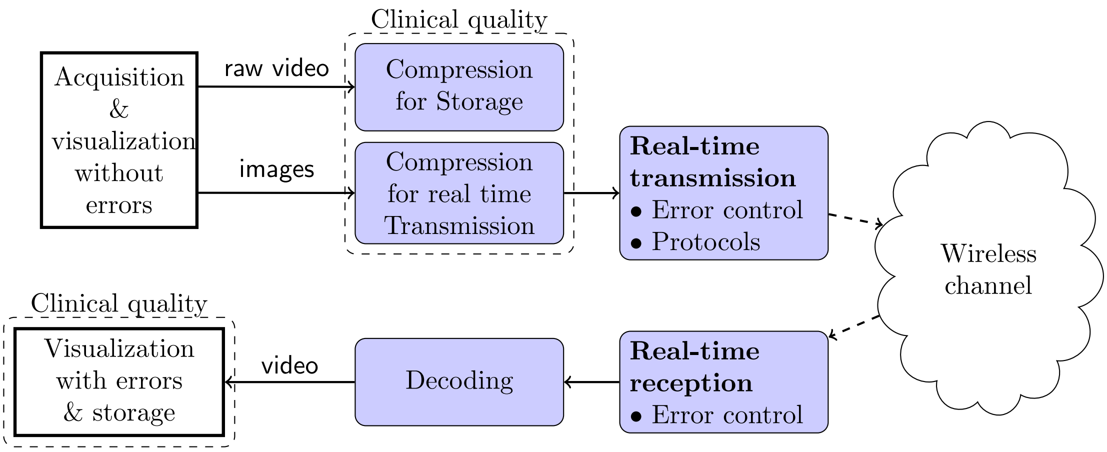


Figure 2.1: *Tele-echocardiography system structure followed in this Thesis.*

imaging storage and transfer over large geographical areas, providing the basis for picture archiving and communications systems (PACS). Moreover, DICOM is the most extensively used standard for storage purpose, recommended by the EAE [12], the ASE [20] and the American College of Cardiology (ACR) [31]. It is also included in the majority of medical image devices [31–33]. For this reason, it is very important to design a compression format interoperable with DICOM and to take into account the DICOM characteristics, described below in Section 2.1.1.

In order to enhance medical image compression while preserving diagnostic information, the concept of Region of Interest (ROI) has been adopted and widely used [34–38]. The challenge of this technique is to perform the image segmentation. In some medical image modalities an algorithm has been designed to obtain the regions automatically. However, obtaining the regions automatically is not always possible, and for almost all medical image modalities the regions of interest for each individual image must be defined by the cardiologist. In order to easily extract the echocardiogram regions, the facilities included in the acquisition devices to segment the images must be taken into account. The devices form the echocardiogram with the different regions: ultrasound image, ECG, auxiliary image and text. A proof that the ultrasound devices incorporate the division of the regions is the calibration DICOM header, which is described in the next Section 2.1.1. In the calibration header different regions are defined with different calibrations that correspond to the ultrasound image region and the ECG. The ROI regions are the ultrasound image and the ECG. The whole ultrasound image has to be selected as ROI. If a small ultrasound part is selected instead, the clinical quality is affected by the degradation quality of the non ROI part [39]. As regards the type of data, two types of region can be distinguished, image and text, since it is not efficient to compress the text as an image. In conclusion, better compression performances can be achieved if the facilities that the ultrasound devices incorporate to divide the image into regions are used to encode each region separately and take into account the data type of each region as well as its clinical importance.

Image compression methods that use wavelet transforms have been successful in providing high compression ratios while maintaining good image quality: Joint Photographic Experts Group (JPEG) 2000 [40] and Set Partitioning in Hierarchical Trees (SPIHT) [41]. However, the DI-

COM standard includes the JPEG 2000 standard, but not SPIHT. Moreover, JPEG 2000 has been demonstrated to achieve very good results in the compression of medical images [42, 43]. It is shown in [43,44] that with a compression rate of 1 bit per pixel (bpp) the diagnostic information is preserved for computerized radiography and ultrasound images, respectively. Consequently, JPEG 2000 can be used for medical image compression providing adequate clinical quality with 1 bpp. The main JPEG 2000 characteristics are described in Section 2.1.2.

2.1.1 DICOM

2.1.1.1 DICOM overview

In response to the increased use of digital images in radiology ACR and the National Electrical Manufacturers Association (NEMA) formed a joint committee in 1983 to create a standard format for storing and transmitting medical images [30]. The committee published the original ACR-NEMA standard in 1985. This has subsequently been revised and in 1993 the standard was renamed DICOM. DICOM is administered by the NEMA Diagnostic Imaging and Therapy Systems division and each year the standard is updated. Details of recent improvements can be found on [30].

The standard describes how to format and exchange medical images and associated information, both within the hospital and also outside the hospital. DICOM interfaces are available for connection of any combination of the following categories of digital imaging devices: (a) image acquisition equipment such as computed tomography, magnetic resonance imaging, computed radiography, ultrasonography, and nuclear medicine scanners; (b) image archives; (c) image processing devices and image display workstations; (d) hard-copy output devices such as photographic transparency film and paper printers.

DICOM addresses five general application areas:

1. Network image management.
2. Network image interpretation management.
3. Network print management.
4. Imaging procedure management.
5. Off-line storage media management.

DICOM is a message standard that facilitates interoperability of medical imaging equipment by specifying:

1. For network communications, a set of protocols to be followed by devices claiming conformance to the standard.
2. The syntax and semantics of Commands and associated information which can be exchanged using these protocols.
3. For media communication, a set of media storage services to be followed by devices claiming conformance to the standard, as well as a File Format and a medical directory structure to facilitate access to the images and related information stored on interchange media.

2.1.1.2 DICOM File Format

A single DICOM file contains both a header (which stores information about the patient's name, the type of scan, image dimensions, etc), as well as all of the image data. The header and the image data are stored in the same file. The image data follows the header information.

Table 2.1: *Some fields of the DICOM header from an Acuson device.*

Field	Contents
Filename	[1x65 char]
FileModDate	“12-nov-2010”
FileSize	2361370
Format	“DICOM”
FormatVersion	3
Width	1024
Height	768
BitDepth	8
ColorType	“truecolor”
FileMetaInformationGroupLength	204
MediaStorageSOPClassUID	“1.2.840.10008.5.1.4.1.1.6.1”
TransferSyntaxUID	“1.2.840.10008.1.2.1”
ImplementationClassUID	“1.2.276.0.7230010.3.0.3.5.4”
Modality	“US”
Manufacturer	“SIEMENS”
InstitutionName	“HC LOZANO BLESA”
ManufacturerModelName	“ACUSON SC2000”
PatientName	[1x1 struct]
PatientID	“XXXXXXXXXXXX”
PatientBirthDate	“XX”
PatientSex	“X”
HeartRate	88
SequenceOfUltrasoundRegions	[1x1 struct]

The size of the header varies depending on the acquisition device and image type. The DICOM elements required depend on the image type that are listed in Part 3 of the DICOM standard [45]. DICOM requires a 128-byte preamble (these 128 bytes are usually all set to zero), followed by the letters 'D', 'I', 'C', 'M'. This is followed by the header information, which is organized in groups: general information, patient, study, series, frame of reference, equipment and image information. In Table 2.1 some fields of a header for an ultrasound device are shown. Of particular importance is the “Transfer Syntax Unique Identification” which reports the structure of the image data, revealing whether the data has been compressed or not. Another important field in the DICOM header included in the ultrasound is the regions calibration, see “SequenceOfUltrasoundRegions” in Table 2.1. It defines regions on the ultrasound image with different calibration and the calibration parameters in order to be able to perform measurements on the ultrasound regions. The calibration header is defined in Part 3 of the DICOM standard [45]. The regions definition depends on the echocardiogram devices and not all the devices define these regions. In Figure 2.2 the calibration regions for an M mode are shown. There are four calibration regions that are defined with four coordinates each one: "Region Location Min X0", "Region Location Min Y0", "Region Location Max X1" and "Region Location Max Y1". The "Region Spatial Format" and the "Region Data Type" of each region indicates the type of mode and data within the region. For example M mode or 2-D mode (tissue or flow) and color bar or spectral (CW or PW Doppler).

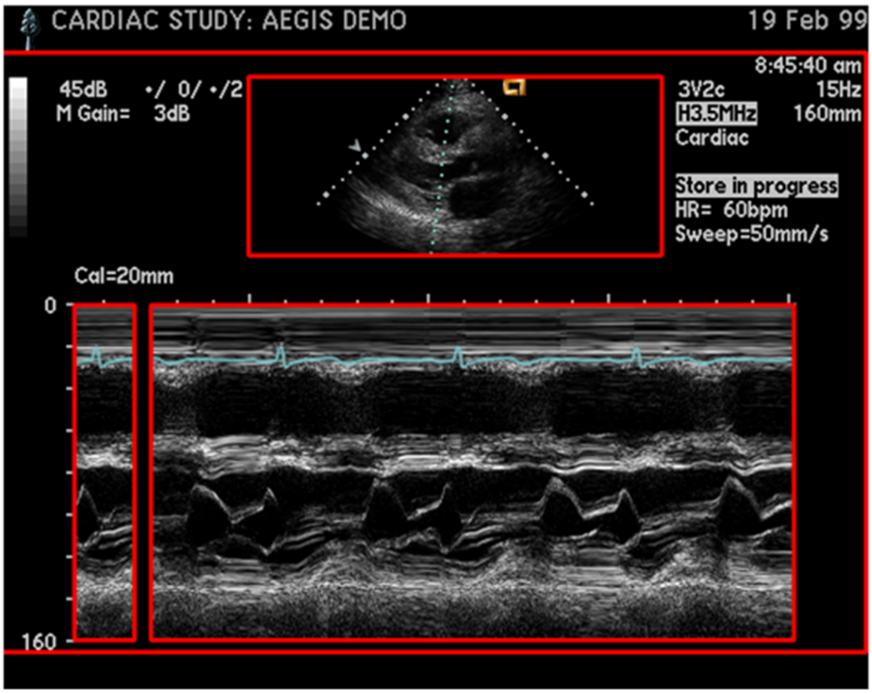


Figure 2.2: *Calibration regions for the M mode of an echocardiogram acquired with an Agilent device.*

The DICOM image exam can be compressed either lossless or lossy in order to reduce disk space. The image format is specified in the "Transfer Syntax Unique Identification" header. The codecs included in DICOM are described in Part 5 of the standard. The image formats supported for DICOM are raw data, lossless Run Length Encoding, JPEG lossy and lossless mode, JPEG-LS [46] lossless and near-lossless mode, JPEG 2000 [40] lossless and lossy mode, MPEG-2 MP@ML and MP@HL image compression, and MPEG-4 AVC/H.264 high profile video compression.

2.1.2 JPEG 2000

JPEG 2000 is an image compression standard that uses the state of the art wavelet technology. It was created by JPEG committee in 2000 with the intention to solve most of the limitations of the original JPEG standard (created in 1992) based on discrete cosine transform.

2.1.2.1 JPEG 2000 Features

JPEG 2000 has many features, which were not available in most of the previous image coding standards. They include:

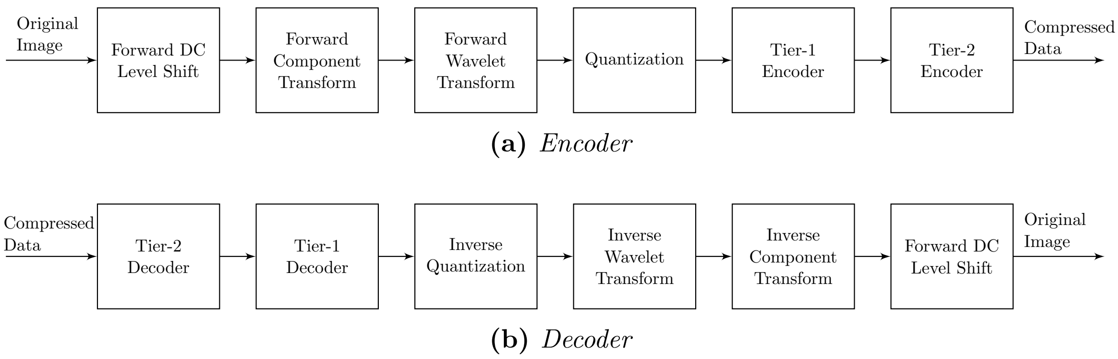


Figure 2.3: *Structure of the JPEG 2000 encoder and decoder.*

- Excellent coding performance. It features superior rate-distortion and subjective image quality performance especially at low bit rates. This is useful in applications whereby file size or transmission time is critical.

- Lossless and lossy compression. It is capable of lossless compression, which is important to some medical imagery and image archival applications.

- ROI coding. It allows certain areas of an image to be encoded at higher fidelity. More information on this feature can be found in [36].

- Spatial and Signal-to-noise ratio (SNR) scalability. It allows progressive recovery of images by resolution or quality.

- Good error resilience. It has added bitstream robustness to the presence of bit errors.

- Flexible file format. The JP2 and JPX file formats allow for handling of color-space information, metadata, and for interactivity in networked applications as developed in the JPEG Part 9 JPEG 2000 Interactive Protocol (JPIP) protocol.

2.1.2.2 JPEG 2000 Encoder and Decoder Structure

As depicted in Figure 2.3a, the core structure of the JPEG 2000 encoder follows a typical sequence of operations used in a transform coding scheme, which consists of transformation, quantization and entropy coding. The JPEG 2000 encoder works as follows. Firstly, the original image with unsigned data is DC level shifted. Then, the component transformation can be carried out if the original image has multiple components. This procedure provides decorrelation among image components and hence improves compression efficiency. There are two component transforms available: one is reversible and may be used for lossy or lossless coding, while the other is irreversible and may only be used for lossy coding. Before proceeding further, it should be noted that the image components can be partitioned into tiles, which are rectangular non overlapping blocks, and thus creating tile components that can be compressed independently of each other.

Wavelet transform may be performed on the tile components. In the lossy case, an irreversible Daubechies 9-tap/7-tap filter is employed, whereas in a lossless case, a reversible 5-tap/3-tap filter is used. The wavelet transform decomposes the tile-components into different decomposition levels, each of which contains a number of subbands filled with transform coefficients. Before entering into the entropy coding phase, the quantization process is carried out to reduce the precision of the transform coefficients. Note that for the lossless case, the quantizer is set to one, i.e. no loss in precision.

The remaining encoding process is grouped into two tiers. In the tier-1 encoder, the quantized transform coefficients associated with each subband are arranged into rectangular blocks called code-blocks. Then, a bit-plane coding technique with three coding passes is applied to each code-block, and the symbols that it produces are coded using an adaptive binary arithmetic coder. In the tier-2 encoder, the inclusion and the order of appearance of bit-plane coding passes along with the actual coding pass data are assembled together to form the final compressed data.

The core structure of the JPEG 2000 decoder is illustrated in Figure 2.3b. The decoding process is basically the reverse of the encoding one.

2.2 Compression for Real-time Transmission

The development of medical video streaming has been made possible thanks to the evolution of efficient video compression techniques. An efficient compression method is extremely important especially for real-time transmission in wireless networks given that they are band limited, time varying, and error-prone. The main objectives in the codification process are:

- Improving the compression method. As a result, less bandwidth is required, reducing the cost, reducing the transmission time and facilitating the arrival of the video at the transmitter without losing diagnostic information.
- Low complexity of the codification process in order to save battery energy and to allow transmission in real-time (the codification time has to be shorter than the time of the image production).

Some of the telemedicine systems dealing with medical video transmission incorporate ROI codification [34, 47–52]. In some medical modalities an algorithm has been designed to obtain the regions automatically, as is the case in [47, 50, 52] for carotid artery ultrasound video. However, obtaining the regions automatically can be a complex process and may introduce excessive delay. For almost all medical image modalities the regions for each individual image must be defined by medical specialists [34, 48, 49, 51]. As already mentioned for the storage encoding, in order to easily extract the echocardiogram regions, the way acquisition devices form the displayed images must be taken into account. Furthermore, each region has different characteristics. For example, it is not efficient to encode text regions as images. Some regions remain invariant for a period of time, i.e. they are images, and as a result it is not efficient to codify these regions as video. Consequently, if the echocardiogram visualization characteristics are taken into account and different codification techniques are applied for each region according to its data type and diagnostic importance, better compression performances can be achieved.

The use of MPEG codecs [53] are widespread due to their high efficiency. The MPEG-4 codec is one of the most commonly used codecs for compressing video sequences. The Xvid codec [54] is an open source implementation of the MPEG-4 standard part 2. The H.264 codec, MPEG-4 Part 10 or AVC [55], was jointly developed by the MPEG and International Telegraph Union Telecommunication Standardization Sector (ITU-T) Video Coding Experts Group (VCEG) who formed the Joint Video Team (JVT). For example, Xvid has been used in [56] and H.264/AVC in [49–52,57] for telemedicine applications dealing with ultrasound video transmission. H.264/AVC has also been used for other medical video modalities [34,58]. The successor of H.264/AVC, High Efficiency Video Coding (HEVC) [59], has been recently incorporated into ultrasound video transmission systems [60,61] showing coding efficiency gains compared to the H.264/AVC. Other very popular compression methods for medical images and video due to their good compression performances are SPIHT for images and 3-D SPIHT [62] for video. SPIHT is based on a 2-D wavelet transform and has features such as embeddedness for progressive transmission, precise rate control for CBR traffic, low complexity and multi-resolution scalability, which are very good qualities for real-time transmission. 3-D SPIHT is an extension of SPIHT from two to three dimensions, two spatial and one temporal, and it shares the SPIHT characteristics. In addition, SPIHT has been demonstrated to provide good results in the compression of ultrasound images [37,63,64], other medical modalities [64–66] and in telemedicine applications dealing with progressive image transmission [67,68]. However, none of the reviewed ultrasound transmission systems implement this compression technique. Consequently, due to their good performance, SPIHT algorithms are proposed for the compression of the echocardiogram in this Thesis: for images without color 2-D SPIHT [41], for images with color 2-D CSPIHT [69] and for video 3-D SPIHT [62]. The main SPIHT characteristics are described below.

2.2.1 SPIHT

The SPIHT algorithm was introduced the first time by Said and Pearlman [41] for the image compression. This algorithm can be seen as an extension of Embedded Zero-tree Wavelet (EZW) introduced by J. M. Shapiro [70]. The EZW technique not only was competitive in performance with the most complex techniques, but was extremely fast in execution and produced an embedded bit stream. With an embedded bit stream, the reception of code bits can be stopped at any point and the data can be decompressed and reconstructed. The encoding algorithms can be stopped at any compressed file size or let run until the compressed file is a representation of a nearly lossless image. In order to have lossless compression specific wavelet transform is necessary. Following that work, in [41] an extension that achieved even better results was developed, SPIHT. EZW was extended to 3-D by Chen and Pearlman [71] and showed promise of an effective and computationally simple video coding system without any motion compensation, and obtained excellent numerical and visual results. 3-D SPIHT algorithm has also been used with excellent results in compression of video [62,72], keeping the simplicity of 2-D SPIHT, while still providing high performance, full embeddedness, and precise rate control. The 2-D SPIHT algorithm and the 3-D extension are shortly described in the following sections.

2.2.1.1 2-D SPIHT

The SPIHT algorithm utilizes three basic concepts:

- Searching for sets in spatial-orientation trees in a wavelet transform.

- Partitioning the wavelet transform coefficients in these trees into sets defined by the level of the highest significant bit in a bit-plane representation of their magnitudes.
- Coding and transmitting bits associated with the highest remaining bit planes first.

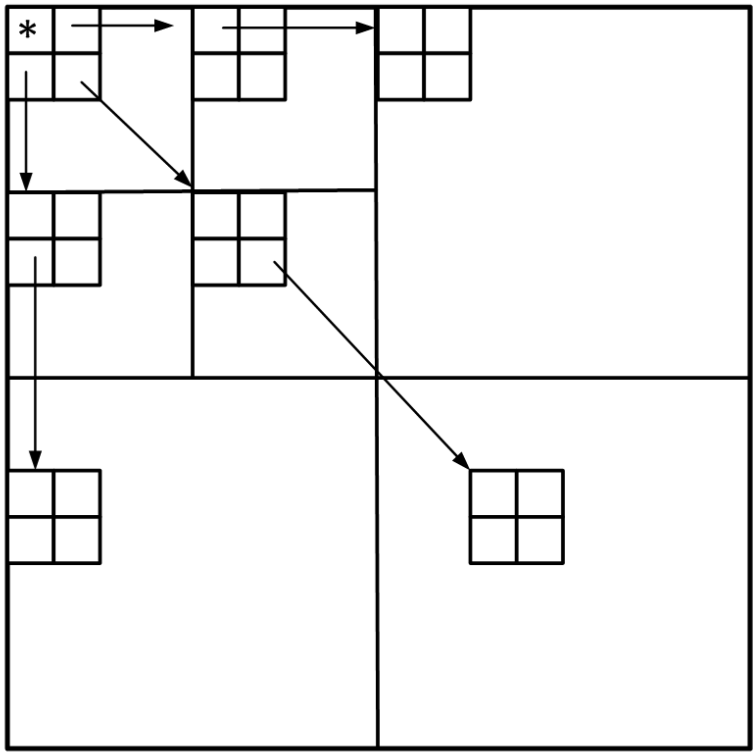


Figure 2.4: *Example of parent-offspring dependencies in the spatial-orientation tree.*

Spatial orientation trees are groups of wavelet transform coefficients organized into trees rooted in the lowest frequency or coarsest scale subband with offspring in several generations along the same spatial orientation in the higher frequency (resolution) subbands. Figure 2.4 depicts the key for parent-offspring relationship of coefficients to tree nodes for a 2-D wavelet transform with two levels of decomposition. In the spatial orientation trees, each node consists of 2 x 2 adjacent pixels, and each pixel in the node has four offspring, except at the highest level of the pyramid, that does not have any offspring. Spatial orientation trees were introduced to exploit self-similarity and magnitude localization properties in a 2-D wavelet-transformed image. In particular, if a coefficient magnitude in a certain node of a spatial orientation tree does not exceed a given threshold, it is very likely that none of its descendants will exceed that threshold.

SPIHT consists of two main stages, sorting and refinement. In the sorting stage, SPIHT sets a magnitude threshold 2^n, where n is called the level of significance, and seeks to identify three entities in the spatial-orientation trees: isolated coefficients significant at level n (magnitude no less than 2^n); isolated coefficients insignificant at level n (magnitude less than 2^n); and sets of coefficients insignificant at level n (all their magnitudes less than 2^n). For a given n, the algorithm searches each tree, partitioning the tree into sets of the three entities above and moves their coordinates respectively to one of three lists: 1) the list of isolated significant pixels (LSP); 2) the list of isolated insignificant pixels (LIP); and 3) the list of insignificant sets (LIS). The last set can be identified by a single coordinate, due to the partitioning rule in the search, where the set of descendants having a significant member is split into its (four) offspring and a subset of all descendants of offspring. When a coefficient is tested and found insignificant, a 0 bit is emitted to the output bit stream and its coordinate is moved to the LIP for subsequent testing at lower n. When a coefficient is found significant, a 1 bit and a sign bit are emitted and its coordinate is moved to the LSP. When an LIS set is tested for significance at level n, a 0 bit is emitted if insignificant. But when found significant, a 1 bit is emitted and the set is partitioned into offspring and descendants of offspring. The offspring are moved to the end of the LIP and subsequently tested for significance at the same n and also to the LIS as roots of their descendant sets that are subsequently tested for significance at the same n.

The bit significance number n is successively lowered in unit increments from the maximum n_{max} of the largest magnitude coefficient. At a given n, the nth of every member of the LSP found significant at a higher n is emitted to the codestream, adding to the $1s$ in the nth bit of the coefficients just found significant for the same n. This is called the refinement stage of the algorithm. When n is decremented, the LIP is tested for significance, and the test result is emitted as a 0 or 1 bit for insignificant or significant, respectively. If significant, its coordinates are moved to the LSP and a sign bit is emitted. Then the LIS is visited and its tree sets are partitioned according to the results of the significance tests. The process terminates when the desired bit rate or quality level is reached.

The decoder of the code bitstream receives the outputs of the significance tests and can therefore build the same lists, the LIP, LIS, and LSP, as in the encoder. Therefore, as input bits are read from the codestream, it reconstructs the magnitude and sign bits of LSP members seen by the encoder. The coefficients of the final LIP and LIS sets are set to zero. In the wavelet transform of an image, large sets of zero values exist which are identified efficiently by SPIHT with a single bit. Moreover, significant coefficients are never represented by more bits than needed in their natural binary representation, since the highest 1 bit is always known.

SPIHT has been originally designed for monochrome images, which are very common in medical applications. A straightforward application to color images is to code the transformed data from each spectral channel independently without exploiting any correlation that might exist between the spectral channels. Alternately, in [73] KahunenLoève Transform (KLT) is performed on the spectral components of the image before coding the decorrelated color planes independently using the SPIHT scheme, called SPIHT KLT. Conversely, the color-EZW (CEZW) scheme [74,75] attempts to exploit underlying spectral correlation by expanding the spatial orientation tree structure used in EZW across the spectral planes for all nodes. However, CEZW has shown to have poorer performance than SPIHT KLT in [74]. In [69], a simple embedded colorcoding scheme based on SPIHT, called color-SPIHT (CSPIHT) showed much better quality of reconstruction without the need to perform a transform of the spectral components of the image. Most recent coding algorithms in which the correlation between the color and Y components is used (such as the Inter Color Correlation based Color SPIHT [76]), are not appropriate because the color components for the ultrasound images are very independent of the Y component.

2.2.1.2 3-D SPIHT

3-D SPIHT is an extension of SPIHT from two to three dimensions and it shares the SPIHT characteristics. In this way, the compressed bit stream will be completely embedded, so that a

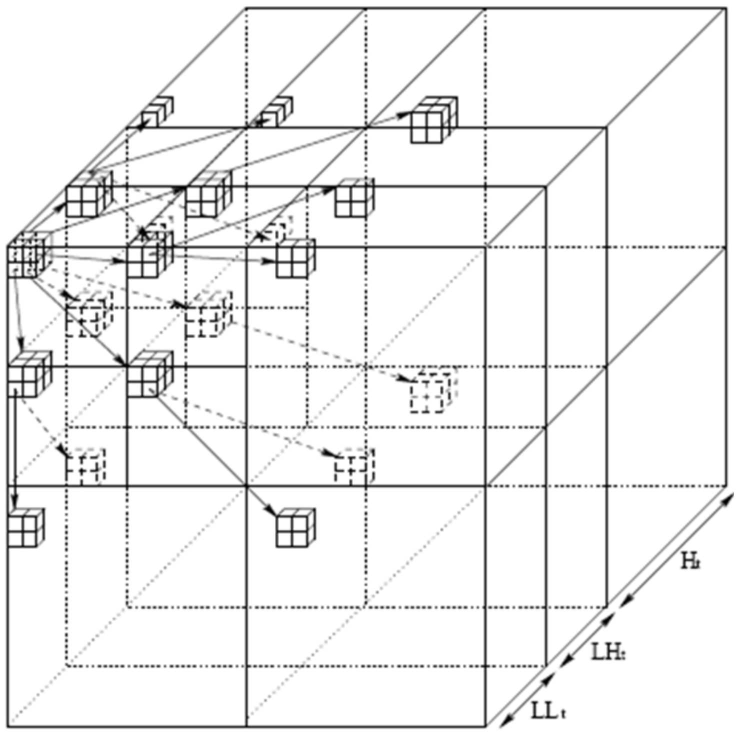


Figure 2.5: *Parent-offspring dependency in 3-D SPIHT at the highest level.*

single file for a video sequence can provide progressive video quality, i.e., the algorithm can be stopped at any compressed file size or let run until nearly lossless reconstruction is obtained.

In 3-D SPIHT, sorting of pixels proceeds just as it would with 2-D SPIHT, the only difference being 3-D rather than 2-D tree sets. Once the sorting is done, the refinement stage of 3-D SPIHT will be exactly the same. On the 3-D subband structure, we define a new 3-D spatiotemporal orientation tree and its parent-offspring relationships. When the spatial and temporal filtering alternate, so that the decomposition is purely dyadic, a straightforward extension from the 2-D case is to form a node in 3-D SPIHT as a block with eight adjacent pixels, two extending to each of the three dimensions, hence forming a node of 2 x 2 x 2 pixels. The root nodes (at the highest level of the pyramid) have one pixel with no descendants and the other seven pointing to eight offspring in a 2 x 2 x 2 cube at corresponding locations at the same level. For nonroot and nonleaf nodes, a pixel has eight offspring in a 2 x 2 x 2 cube one level below in the pyramid. Figure 2.5 depicts these parent-offspring relationships in the case of a two-level dyadic 3-D decomposition with 15 subbands, produced by a once repeated spatial-horizontal, spatial-vertical, and temporal splitting, in that order.

The number of temporal subband decomposition directly depends on the number of frames to be compressed at a time. The typical value for the temporal resolution is of sixteen frames, although a shorter filter with four or eight frames, such as the Haar or S+P filters, can be also used. Only two or three decompositions are possible with four or eight frames compressed at a time, respectively. However, with a spatial resolution of 352 x 288 pixels, for example, five spatial (dyadic) decompositions can be achieved with the high performance 9/7 biorthogonal filter [77].

The next step is compression of the coefficients into a bit stream. Essentially, it can be done by feeding the 3-D data structure to the 3-D SPIHT coding kernel. The 3-D SPIHT kernel will sort the data according to the magnitude along the spatio-temporal orientation trees (sorting pass), and refine the bit plane by adding necessary bits (refinement pass). At the destination, the decoder will follow the same execution path conveyed by the received significance decision bits to recover the data.

YYYYYYYYYYYYYYYY	UUUU	UUUU

(a)

YYYYYYYYYYYYYYYY UUUUUUUUUUUUUUUU VVVVVVVVVVVVVVVV

(b)

Figure 2.6: *Bit stream of two different methods: (a) separate color coding and (b) embedded color coding.*

A simple application of the SPIHT to color video would be to code each color plane separately, as does a conventional color video coder. Then, the generated bit stream of each plane would be serially concatenated. However, this simple method would require allocation of bits among color components, losing precise rate control, and would fail to meet the requirement of the full embeddedness of the video codec since the decoder needs to wait until the full bit stream arrives to reconstruct and display. Instead, one can treat all color planes as one unit at the coding stage, and generate one mixed bit stream so that we can stop at any point of the bit stream and reconstruct the color video of the best quality at the given bit rate. In addition, we want the algorithm to automatically allocate bits optimally among the color planes. By doing so, we will still keep the claimed full embeddedness and precise rate control of 3-D SPIHT. The bit streams generated by both methods are depicted in the Figure 2.6, where the first one shows a conventional color bit stream, while the second shows how the color embedded bit stream is generated, from which it is clear that we can stop at any point of the bit stream, and can still reconstruct a color video at that bit rate as opposed to the first case.

Let us consider a tri-stimulus color space with luminance Y plane such as YUV or YCrCb. Each such color plane will be separately wavelet transformed, having its own pyramid structure. Now, to code all color planes together, the 3-D SPIHT algorithm will initialize the LIP and LIS with the appropriate coordinates of the top level in all three planes. Since each color plane has its own spatial orientation trees, which are mutually exclusive and exhaustive among the color planes, it automatically assigns the bits among the planes according to the significance of the magnitudes of their own coordinates. The effect of the order in which the root pixels of each color plane are initialized will be negligible except when coding at extremely low bit rate. Note also that the wavelet transforms and sizes may be different among the three planes without affecting the method.

In conclusion, 3-D SPIHT is an embedded subband-based video coder analogous to the 2-D spatial orientation trees in image coding. The video coder is fully embedded, so that different degrees of monochrome or color video quality can thus be obtained with a single compressed bit stream. The cost for this embeddedness is the coding delay (latency) to accept 16 frames into a buffer and a memory size of the order of the size of the coding unit to execute the 3-D SPIHT algorithm. Precise rate control and self-adjusting rate allocations are automatically achieved. In addition, spatial and temporal scalability can be easily incorporated into the system to meet various types of display parameters requirements.

2.3 Transmission Protocols

In order to transmit the coded information over Internet networks, Transmission Control Protocol (TCP)/Internet Protocol (IP) is the suite of communications protocols used to connect hosts on the Internet. TCP/IP provides end-to-end connectivity specifying how data should be formatted, addressed, transmitted, routed and received at the destination. It has four abstraction layers (see Figure 2.7) which are used to sort all related protocols according to the scope of networking involved:

Application
Transport
Internet
Link

Figure 2.7: *TCP/IP layer protocol stack.*

- The link layer contains communication technologies for a single network segment of a local area network. The link layer depends on the physical medium used for the transmission.
- The Internet layer has the responsibility of sending packets across potentially multiple networks. Internet working requires sending data from the source network to the destination network. This process is called routing. In the Internet protocol suite, the Internet Protocol performs two basic functions:
  - Host addressing and identification: this is accomplished with a hierarchical IP addressing system.
  - Packet routing: this is the basic task of sending packets of data (datagrams) from source to destination by forwarding them to the next network router closer to the final destination.
- The transport layer protocol establishes host-to-host connectivity, meaning it handles the details of data transmission that are independent of the structure of user data and the logistics of exchanging information for any particular specific purpose. Its responsibility includes end-to-end message transfer independent of the underlying network, along with error control, segmentation, flow control, congestion control, and application addressing (port numbers). End to end message transmission or connecting applications at the transport layer can be categorized as either connection-oriented, implemented in TCP, or connectionless, implemented in User Datagram Protocol (UDP).

- The application layer contains the higher-level protocols used by most applications for network communication. Data coded according to application layer protocols are then encapsulated into one or (occasionally) more transport layer protocols, which in turn use lower layer protocols to effect actual data transfer.

Transport protocols are further distinguished in upper and lower layer, Real-Time Transport Protocol (RTP), and UDP/TCP respectively. TCP [78] provides reliable, ordered, error-checked to secure packet delivery to destination. These characteristics may lead to long delay and jitter, specially when high data flows are transmitted. Thus, TCP is highly adequate for applications where it is important to guarantee the reliability while the delay is less important. On the other hand, UDP [79] does not provide any error handling or congestion control mechanisms, allowing therefore packets to drop out. This properties made UDP highly adequate for applications where the packets arrive in time is more important than the integrity of the information. Therefore, UDP is the primarily established as the lower layer transport protocol for real-time transmission of video. However, as a minimal clinical quality is necessary to achieve an adequate diagnostic error control methods should be implemented in the upper layer.

RTP [80] is widely used in clinical video steaming applications [49, 50, 57]. RTP provides end-to-end delivery services for real-time video and audio transmission. RTP itself does not contain any mechanisms to ensure on time delivery. On the contrary, it relies on UDP or TCP for doing so. However, it does provide the appropriate functionality for carrying real-time content such as time-stamping and control mechanisms that enable synchronization of different streams with timing properties.

Since the echocardiogram is composed of several regions with different type of data (text, image, video, audio) and different clinical importance and consequently every region can be compressed with different coding methods, different transmission methods can be applied for each region. RTP is not suitable for this proposal, since it does not provide delivery of text or images. Given the aforementioned, a protocol for end-to-end real-time transmission of echocardiograms over IP can be designed introducing different coding and transmission protocols for every region, UDP or TCP and further control error methods. The most suitable protocols will depend on the type of region and its clinical information. For example, in order to transmit a region with the ultrasound video, UDP is the adequate protocol introducing additionally an error control method in the application layer. However, in order to transmit the text present in the echocardiogram, TCP is more adequate, because the reliability of the transmission is required.

In order to decrease header overheads, reduce packet loss and increase security over noisy wireless links, RObust Header Compression (ROHC) can be used for both UDP [81] and TCP [82] transport layers. This standard compresses IP and UDP headers to just 3 bytes, including the checksum field to discard erroneous packets, and the IP and TCP headers to just 10 bytes. Thanks to this standard, the redundancy introduced in the transmission is decreased, allowing small packets to be transmitted. This reduces packet loss without an excessive increase in the number of transmitted bits.

It is important to take into account that some text regions may contain confidential information about the patient. For this reason, the text has to be protected so that it can only be accessed by authorized sanitary staff. An easy way to protect this information is by protecting all the packets in which the information is contained. Typically, the text information is not protected in the clinical video streaming systems since the text is included in the video. Transport Layer Security (TLS) [83] is the most common choice for secure communications and is included in the medical standards DICOM and Health Level Seven (HL7). The primary advantage of TLS is that it provides a transparent connection-oriented channel. Thus, it is easy to secure an application protocol by inserting TLS between the application layer and the transport layer.

2.4 Error Control Methods

In order to provide an accurate diagnosis, it is not only necessary to have a compression method that guarantees clinical quality, but it is also essential to be able to guarantee the integrity of the video during the transmission process. It is well known that wireless channels are error prone. Thus, digital transmission may be affected by erroneous bits that distort the reconstruction of the video on reception. Hence, the use of error control mechanisms for maintaining acceptable video levels in wireless communications channels are required [84]. However, for real-time applications dealing with multimedia data, reliable methods such as the incorporated in TCP are not recommended due to the resulting delay. Error control methods can be used in the application layer for the ROI regions in order to minimize distortion, but without introducing excessive delay.

The two main error control techniques are Automatic Repeat ReQuest (ARQ) [85–87] and Forward Error Correction (FEC) [88–92]. Both techniques increase the original amount of bits and therefore there is a trade off between compression fidelity and protection. With retransmission techniques, only missing packets are retransmitted. If the channel delay is long or if several retransmissions of the same packet are required because of channel errors, the resulting delay would be intolerable for real-time applications. With FEC, redundant bits are added in the transmitted data and the decoder uses these added bits to correct the errors. The amount of redundancy embedded can be more than is necessary to correct the errors, using more bits than with the retransmission mechanism. However, if the amount of redundancy is less than is necessary, no error can be corrected. But this technique does not need a feedback channel and reduces the time needed to recover missing packets. The design and performance of a hybrid ARQ with concatenated FEC for real-time video streaming over wireless networks has been addressed in [93, 94]. In [93], an adaptive technique in the Medium Access Control (MAC) layer of Worldwide Interoperability For Microwave Access (WiMAX) was used, but this technique is only valid for the WiMAX channel. In [94] a more general solution was proposed. Furthermore, the channel conditions can be taken into account to adapt the error control techniques to the channel conditions and use one or another techniques.

For telemedicine applications, error resilient implementations for robust diagnosis performance have been addressed, for example, in [34, 49–52, 57, 95]. A scalable video coding (SCV) employing spatiotemporal scalability is found in [34, 57]. A flexible macroblock ordering (FMO) technique for variable quality slice encoding and redundant slices for resilience over error prone mediums, available in the baseline profile of H.264/AVC, were used in [50, 52]. In [49] a ROI-based and prediction-based unequal channel error protection implementation to overcome transmission errors was presented. Another different approach is an optimized cross-layer design (CLD) based on a reinforcement learning algorithm for real-time medical video streaming [51, 95]. In [51] an error

Table 2.2: *Mobility, data rate and delay of 3G and beyond wireless access technologies.*

Wireless technology		Max. speed	Data rate[1]	delay
3G	UMTS [96]	300 km/h	220 - 384 Kbps	<250 ms
3.5G	HSPA [97,98]	300 km/h	500 kbps - 2 Mbps	<250 ms
	HSPA+ [99,100]	300 km/h	1 - 4 Mbps	<100 ms
	LTE [100,101]	500 km/h	1.5 - 5.8 Mbps	<70 ms
4G	LTE-Advanced [102]	500 km/h	up to 100 Mbps	<70 ms
WiMAX	IEEE 802.16e [103]	120 Km/h	up to 5.6 Mbps	<70 ms
	IEEE 802.16m [104]	350 Km/h	up to 100 Mbps	<70 ms

[1] Uplink data rates

concealment technique for the ROI region was used. This last technique is not suitable for real time applications. However, a hybrid ARQ with a concatenated FEC method in which the channel conditions are taken into account has not been used in any of the reviewed telemedicine applications for clinical video streaming.

2.5 Wireless Transmission Technologies

The development of more demanding telemedicine systems has been possible thanks to the evolution of mobile telecommunication systems. The evolution of mobile systems from 2 Generation (G) to 2.5G and afterwards to 3G facilitates the provision of higher data rates and lower delays that enable the development of more responsive telemedicine systems. Furthermore, the later wireless technologies, 3.5G and beyond, enable the transmission of high quality video resolution that requires more bandwidth.

Before starting to design a telemedicine system, it is important to know the application requirements such as required bandwidth, delay, and speed in order to choose the appropriate access technology. Table 2.2 shows the main characteristics of the access technologies for 3G and beyond, including maximum speed, typical up link data rates and typical delays. One of the most frequently used technologies in clinical video streaming applications is WiMAX [49, 51, 52, 95] due to its improved uplink and downlink rates, increased coverage and throughput, mobility support enhancement, latencies reduction, quality of service (QoS) services and security enhancement.

2.5.1 Worldwide Interoperability For Microwave Access (WiMAX)

The Institute of Electrical and Electronics Engineers (IEEE) 802.16 standard offers broadband wireless access (BWA) over long distance. WiMAX was firstly standardized for fixed wireless access by the IEEE 802.16-2004 [105] and then for mobile access with by the IEEE 802.16e [103] and 802.16m [104] standards. WiMAX is a promising technology to provide wireless services requiring high-rate transmission and strict QoS requirements in both indoor and outdoor environments. A thorough overview of WiMAX standardization process and evolving concepts and technologies up to IEEE 802.16e standards appears in [106], while recent advances are described in detail in [107–109].

To provide flexibility for different applications, the standard supports two major deployment scenarios:

- Last-mile BWA: In this scenario, broadband wireless connectivity is provided to home and business users in a wireless metropolitan area networks (WMAN) environment. The operation is based on a point-to-multipoint single hop transmission between a single base station (BS) and multiple subscriber stations (SSs).
- Backhaul networks: This is a multihop (or mesh) scenario where a WiMAX network works as a backhaul for cellular networks to transport data/voice traffic from the cellular edge to the core network (Internet) through meshing among IEEE 802.16/WiMAX SSs.

The WiMAX Forum, which is a nonprofit organization, encourages and supports IEEE 802.16-based BWA. The main role of the WiMAX Forum is to standardize and maintain the process of testing and the certification program for compatibility and interoperability of IEEE 802.16 equipment [110]. The physical (PHY) and MAC layer protocols are well defined in the IEEE 802.16 standard, efficient radio resource management is still an open issue. Features of PHY and MAC layers are discussed next.

2.5.1.1 Physical Layer Features

The physical layer of the IEEE 802.16 air interface originally operates at either the 1066 GHz (IEEE 802.16), 211 GHz band (IEEE 802.16a) [111]. Todays licensed deployment is typically in the range of 2.3, 2.5-2.7, 3.5, and 5.8 GHz, while 4G frequency bands will facilitate deployment between 450-3600 MHz [104]. Channel bandwidth allows great flexibility in the sense that it allows WiMAX operators to consider channel bandwidths between 1.25, 2.5, 5, 10, and 20 MHz (802.16e). In 802.16m scalable bandwidth between 5-40 MHz for a single radio frecuency carrier is considered, extended to 100 MHz with carrier aggregation to meet International Mobile Telecommunications-Advanced (IMT-Advanced) requirements. WiMAX employs a set of high and low level technologies to provide robust performance in both line-of-sight and non-line-of-site conditions.

The primary features of the physical layer include adaptive modulation and coding (QPSK, 16-QAM, 64-QAM), hybrid ARQ, and fast channel feedback. WiMAX uses scalable orthogonal frequency division multiple access (SOFDMA) that divides the transmission bandwidth into multiple subcarriers. The number of subcarriers ranges from 128 for 1.25 MHz channel bandwidth and extends up to 2048 for 20 MHz channels. In this manner, dynamic QoS can be tailored to an individual application requirements. In addition, orthogonality among subcarriers allows overlapping leading to flat fading. In other words, multipath interference is addressed by employing orthogonal frequency-division multiplexing (OFDM) while available bandwidth can be split and assigned to several requested parallel applications for improved systems efficiency. The latter is true for both downlink and uplink. A Multiple input multiple output antenna system improves communication performance, including significant increases in data throughput and link range, without additional bandwidth or increased transmit power.

2.5.1.2 Medium Access Control Layer Features

IEEE 802.16/WiMAX uses a connection-oriented MAC protocol, which provides a mechanism for the SSs to request bandwidth from the BS. Although each SS has a standard 48-bit MAC address, the main purpose of this address is for hardware identification. Therefore, a 16-bit connection identifier is used primarily to identify each connection to the BS. IEEE 802.16/WiMAX supports both frequency division duplex (FDD) and time-division duplex (TDD) transmission modes.

The QoS scheduling is the most important feature in WiMAX systems, which makes it an ideal choice for QoS sensitive applications such as video content streaming. There are three major types of services supported with different QoS requirements:

- Unsolicited grant service (UGS): This service supports CBR traffic. In this case the BS allocates a fixed amount of bandwidth to each of the connections in a static manner. UGS service is suitable for traffic with very strict QoS constraints for which delay and loss need to be minimized. A typical application is VoIP.
- Polling service (PS): This service supports traffic for which some level of QoS guarantee is required. The amount of bandwidth required for this type of service is determined dynamically based on the required QoS performance and the dynamic traffic arrivals for the corresponding connections. It can be divided into two subtypes:
  - Real-time polling service (rtPS). This service is delay sensitive. Typical applications are audio and video streaming.
  - Non-real time polling service (nrtPS). This service can guarantee a certain throughput guarantee. A typical application is file transfer.
- Best-Effort (BE) Service: This is for traffic with no QoS guarantee. The amount of bandwidth allocated to BE service depends on the bandwidth allocation policies for the other two types of service. In particular, the bandwidth left after serving UGS and PS traffic is allocated to BE service. Typical applications are web and email traffic.

Mobility management is also address in 802.16e and current 802.16m standards. Established connections can move with speeds between 50-100 km/h for 802.11e and up to 350 km/h for 802.11m with adequate performance.

2.6 Clinical Quality

The image quality in telemedicine systems is decreased by lossy compression and due to errors introduced by the network. In medical images, a minimum acceptable quality is required to be able to make an adequate diagnosis. In order to quantify the clinical distortion introduced in the transmitted signal, distortion indices should be used. There are two types of indices: objective or mathematical distortion indices and subjective or clinical distortion indices.

2.6.1 Mathematical Distortion Indices

Mathematical distortion indices [112], thanks to their ease of use, can be useful for preliminary and fast measurement of quality. Some examples of these indices are: Mean Squared Error (MSE), Peak Signal-to Noise Ratio (PSNR) and SIMilarity Index [113]. Currently, the most commonly used objective image and video distortion metric even for medical images is PSNR [49, 50, 57]. PSNR is widely used because it is simple to calculate, has clear physical meanings, and is mathematically easy to deal with for optimization purposes. PSNR measured in decibels (dB) is given by

$$PSNR \;=\; 10 \cdot \log_{10}\left(\frac{255^2}{MSE}\right) \tag{2.1}$$

where 255 is the maximum possible value that can be attained by the image signal of 8 bits. MSE is defined as

$$MSE \;=\; \frac{1}{M \cdot N} \cdot \sum_{m=1}^{M} \sum_{n=1}^{N} \left| I\left(x, y\right) - J\left(x, y\right)\right|^2 \tag{2.2}$$

where $M \cdot N$ is the frame dimension in pixels. $I(x, y)$ is the source frame and $J(x, y)$ is the compressed frame.

However, PSNR has been widely criticized for not correlating well with perceived quality measurement [114, 115] and for not being able to quantify the real clinical distortion [56].

2.6.2 Clinical Distortion Indices

New indices that include degradation in the diagnosis content of echocardiograms are necessary. There are two main types of test that can be included to calculate a clinical distortion index:

- Semi-blind test: comparison of the compressed images with the original, without knowing the transmission rate used for the compressed images. Can the same diagnosis be made with both images?
- Blind-test: evaluation of the compressed images without knowing the transmission rate used. In general, for medical images this evaluation consists of performing a complete diagnosis and comparing it to the results of the original images.

There are several works in which medical opinion has been utilized to assess real clinical quality after codification [48, 50, 56] and transmission [17, 18, 48, 50, 57, 116]. However, these assessments were not as accurate as the evaluation in [56]. In [56], a testbed that unifies and reflects clinical evaluation for echocardiogram procedures was developed using both a blind and a semi-blind test. This testbed gives a precise evaluation of the real degradation in compressed echocardiograms and provides a minimum recommended transmission rate required for each echocardiographic mode to achieve adequate clinical quality. However, it has a serious disadvantage in that it is burdensome and very costly in terms of time dedicated to evaluating the echocardiograms. Furthermore, if a guarantee of clinical quality for the visualized echocardiogram is desired, it will be necessary to evaluate a wide range of transmission rates and network scenarios with different channel conditions. In order to save time in the evaluation process, two different evaluations can be performed. In the first evaluation, a wide range of transmission rates can be evaluated with a method as accurate as that presented in [56], but saving time in the evaluation method. In order to achieve this, a two phase testbed can be designed. In the first phase, a fast and simple evaluation can be made using a semi-blind test for different transmission rates in order to select two transmission rates for each mode that are at the limits of acceptable clinical quality. These two rates can be evaluated in more detail in the second phase using a blind test in order to provide the recommended transmission rates. If the transmission occurs without errors and the whole echocardiogram is displayed at the recommended transmission rate, the same diagnosis as that of the original echocardiogram is guaranteed. However, if the recommended transmission rate is not received all the time, the same diagnoses may or may not be possible. Instead of assessing different channel conditions, as has been described in the literature, an evaluation in which the percentage of time that the echocardiogram is displayed with a lower transmission rate than the recommended rate can be evaluated. In this manner, it is possible to know for any channel if the diagnosis on reception will be possible with a single evaluation. As the starting point of the evaluation is the recommended transmission rate in the previous evaluation, for the second evaluation it may be sufficient to conduct a semi-blind test for different percentages of time that the echocardiogram is displayed with a transmission rate inferior to the minimum.

2.7 Tele Ultrasound Systems Overview

Overviews of m-health systems and more recent wireless medical video transmission systems were addressed in [26, 28], respectively. This section discusses the most relevant studies dealing with ultrasound video transmission over wireless channels, the main characteristics of which have been addressed throughout this chapter. Table 2.3 shows the most important parameters for the reviewed systems in order to compare the actual systems to each other and with respect to the system proposed in this Thesis. The most important parameters in a clinical video transmission over a wireless channel are the spatial and temporal (frame rate) resolution, the encoding and error resilient method, the transmission rate, clinical evaluation and the access channel.

The video resolution affects the diagnostic capacity of the video and the transmission rate. The higher the resolution, the higher the quality and the transmission rate [52]. The resolution can be classified as low video resolution (lower than 480x480 and 15 fps) and original/high video resolution (equal to or higher than 480x480 and 15 fps). In Table 2.3, the systems are divided into low (light gray) and high (dark gray) resolution. Higher video resolution requires higher transmission rates, and consequently wireless technologies with high data transfer rates, 3.5G and beyond. Other factors that affect the transmission rate are the encoding and error resilient methods used, thus it is very important to choose these methods correctly. The transmission rate has to be the minimum possible but without compromising the clinical quality of the video. For this reason, it is very important to perform an evaluation that defines the minimum transmission rate for the selected parameters in order to guarantee clinical quality, as in [56], where the lowest transmission rate is required for high resolution videos, see Table 2.3.

Table 2.3: *Ultrasound video transmission systems summary.*

Reference	Resolution	Encoding	Transmission rate	Clinical evaluation	Channel
Pedersen 2009 [57]	320x240, 10 fps	H.264/AVC (Scalable)	349 kbps	Yes	3G and beyond
Martini 2010 [49]	480x256, 15 fps	H.264 unequal error protection based on ROI	300 kbps	No	WiMAX
Panayides 2011 [50]	352x288, 15 fps	H.264/AVC FMO ROI RS	411 kbps	Yes	3G and beyond
Alinejad 2012 [95]	352x288, 20 fps	Windows Media Cross-layer	300 - 500 kbps	Yes	3G and WiMAX
Alesanco 2009 [56]	720x576, 25 fps	Xvid	256 kbps, 756 kbps	Yes	-
Debono 2012 [51]	640x480, 25 fps	H.264/AVC cross-layer Error concealment on ROI	860.16 - 1933.90 kbps	No	WiMAX
Panayides 2012 [117]	704x576, 15 fps	HEVC	1.5 Mbps	Yes	WiMAX
Panayides 2013 [52]	704x576, 15 fps	H.264/AVC FMO ROI RS	1.3 - 1.5 Mbps	Yes	WiMAX

2.8 Conclusions: Thesis approach

This section presents the conclusions of the Chapter: the main improvements that can be made in the different parts of the tele-echocardiography system. This improvements are the design premises used in this Thesis to improve tele-echocardiography systems. They are summarized below for each part of the system.

Compression for storage

- Designing a compression format interoperable with DICOM and taking into account the DICOM format.
- Taking advantage of the segmentation facilities incorporated in the acquisition devices to segment the image into regions and compress each region according to its data type (image or text) and diagnostic importance.
- Compressing the image regions with JPEG 2000 and recommended quality of 1 bpp for the ROI regions so as to guarantee clinical quality.

Compression for real-time transmission:

- Taking advantage of the visualization characteristics of the echocardiograms and the segmentation facilities incorporated in the acquisition devices to compress each region according to its data type (video, image, text, signal or sound) and diagnostic importance.
- Compressing the image and video regions with SPIHT algorithms: for images without color 2-D SPIHT [41], for images with color 2-D CSPIHT [69] and for video 3-D SPIHT [62].

Transmission protocols:

- Designing a protocol for end-to-end real-time transmission of echocardiograms over IP where different transmission protocols are applied for every region, UDP or TCP, and error resilient methods are applied according to the clinical importance of the region in question.
- The compression of headers, ROHC, can be applied so as to reduce the overhead, reduce packet loss and increase security over noisy wireless links.
- The TLS protocol can be applied to the text regions in order to protect patients' confidential information.

Error control method:

Designing an error control method for the application layer and the regions with relevant clinical information that does not use a reliable protocol in the upper layers. ARQ and FEC techniques can be used depending on the channel conditions.

Clinical evaluation:

- Designing an evaluation methodology that consists of two evaluations so as to reduce time in the whole process.
- A first evaluation of the transmission rates for each echocardiogram mode to recommend a minimum transmission rate that guarantees the same diagnosis as that of the original images. This test is divided into two phases in order to reduce the assessment time. The first phase is a semi-blind test to discriminate two transmission rates. The second phase is a blind test for the two transmission rates that has been selected in the previous phase.
- A second evaluation of the percentage of time that the echocardiogram can be visualized with a transmission rate inferior to the recommended rate. This evaluation consists of a semi-blind test. It is possible to know if the images have been displayed with adequate quality without the need to perform an assessment for each channel condition.

Chapter 3

Echocardiogram Compression

This Chapter deals with the first part of tele-echocardiography systems (see Figure 3.1): compression. Compression performance is a key factor for achieving good results in transmission and for saving storage space. In order to design efficient compression methods, the characteristics of the echocardiogram have to be taken into account. Since echocardiograms performed for real-time transmission have different characteristics than those performed for storage purposes, different compression approaches are proposed for each. It is necessary not only to design efficient compression methods but also to guarantee clinical quality. Thus, an evaluation of transmission rates is needed in order to recommend a minimum transmission rate that guarantees the same diagnosis as that of the original echocardiogram. This Chapter is organized as follows. The echocardiogram databases and characteristics for both storage and real-time transmission purposes are described in Section 3.1. Section 3.2 describes the clinical evaluation methodology. Section 3.3 defines the compression approach for storage purposes and Section 3.4 provides the compression results compared with other techniques. Section 3.5 defines the compression approach for real-time transmission purposes and Section 3.6 sets out compression recommendations and results. The results are compared with the results of previous systems. Finally, the conclusions of this Chapter are given in Section 3.7.

3.1 Echocardiogram Databases and Characteristics

An in-depth analysis of echocardiogram characteristics is required in order to design efficient compression and transmission methods. Two types of echocardiograms can be distinguished according to whether the echocardiogram is stored or if it is transmitted in real-time:

- Stored echocardiograms: a clinical compression is applied before the echocardiogram is stored. This avoids having to store the entire echocardiogram video. The stored echocardiogram consists of several images of different heart visualizations, and several video loops of the same heart visualization in movement from 1 to 3 cardiac cycles (from 14 to 64 frames).
- Real-time echocardiograms: no clinical compression is applied. The real-time echocardiogram consists of the whole echocardiogram video.

The composition of both types of echocardiograms is defined by the EAE in [12], and a summary is shown in Table 3.1. The characteristics of the stored echocardiogram depend on the patient's

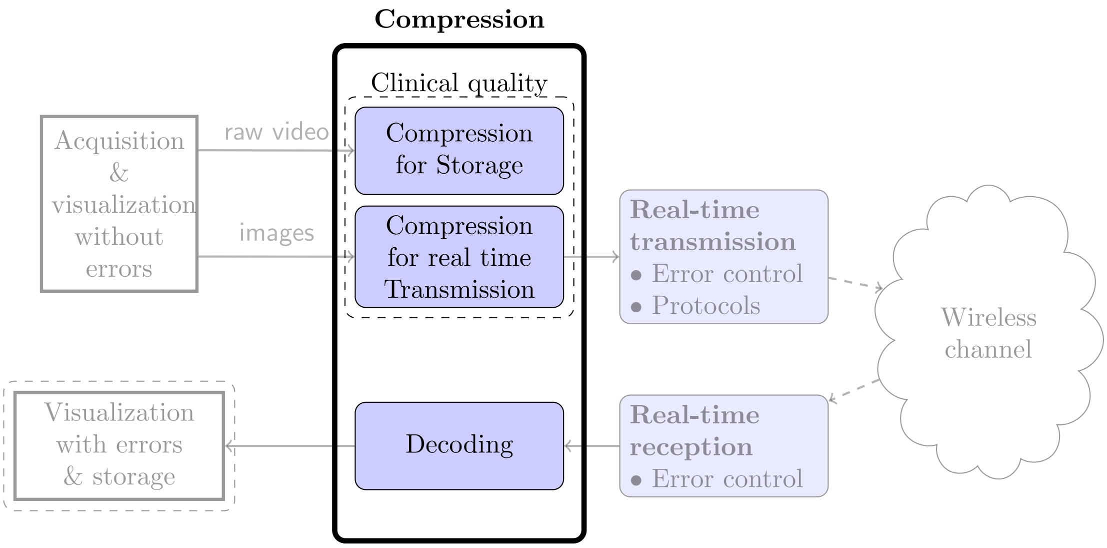


Figure 3.1: *Tele-echocardiography system structure followed in this Thesis: compression part.*

Table 3.1: *Composition of the two types of echocardiogram according to [12].*

Type	No. Images	No. Videos (duration)
Stored	≤ 6	≤ 13 (≤ 3 seconds)
Real-time	-	1 ($\cong 30$ minutes)

condition. The expert cardiologist may decide to store all the views (see Table 3.1) or only those views that present abnormalities.

The echocardiograms have different modes of operation. Each mode presents different representation of the heart. The main modes of operation has been described in Chapter 1. The echocardiogram can be divided into regions according to how the echocardiogram devices form the displayed images, the data type of each region and the clinical relevance of the region. Nevertheless, the data type of each region may be different for both types of echocardiograms. In order to study the echocardiogram characteristics, a database was collected for the two types. Moreover, these databases have been used to carry out different assessments during this Thesis. In the following sections the databases and the characteristics for the two types of echocardiogram are described.

3.1.1 Stored Echocardiograms Database and Characteristics

3.1.1.1 Database for Stored Echocardiograms

In order to analyze the stored echocardiogram characteristics and to evaluate the compression performance for the stored echocardiogram, we have echocardiograms available in DICOM format that correspond to four different devices and different operation modes since each operation mode has different visualization characteristics. These echocardiograms were captured and stored by different cardiologist and from different patients. Table 3.2 shows the echocardiogram devices, the number of images available for each device, the typical image size of each device and the mean DICOM header size for each device. These devices are representative of the typical image distribution in echocardiogram images. The echocardiograms acquired with the Agilent and Siemens Acuson devices can be considered as one because both have a similar size and similar characteristics.

Table 3.2: *Database for the stored echocardiograms. Echocardiogram devices, number of echocardiogram images available, typical image size and mean DICOM header size in bytes of all the exams for each device.*

Device	Number of images	Typical Image Size	DICOM Header Size (bytes)
Agilent Sonos	22	600x430	10365
Siemens Acuson	23	576x456	19132
Philips Envisor	28	800x564	19132
Siemens CS2000	32	1024x768	19132

3.1.1.2 Stored Echocardiogram Characteristics

There are three types of regions, as we can see in Figure 3.2:

- Ultrasound: it is the most important because it contains the relevant information for the diagnosis. The ultrasound is always present and only appears once. In the case of having a video loop, it is the only region that changes over the time.
- Auxiliary images: they surround and complement the ultrasound region. These auxiliary images are, for example, the ECG, the color label, other ultrasound images to complement the information of the main ultrasound image and some symbols regarding the configuration.
- Text: it is always present in all the images and contains information such as patient information, date, time, configuration or measurements of the study.

In the case of having a video loop, the only region that change over time is the ultrasound region, while the rest of the regions remain invariant. For each echocardiogram mode and device, the distribution and the size of the regions, the number of auxiliary regions and the text are different. Furthermore, some image regions contain color information that is relevant for the diagnosis and others not. For example, the Doppler modes have relevant information in color in the ultrasound image and in the color scale.

3.1.2 Real-time Echocardiograms Database and Characteristics

3.1.2.1 Database for Real-Time Echocardiograms

Three cardiologists experienced in echocardiography recorded and stored original echocardiograms from patients with different diagnoses and with three different ultrasound devices (see Table

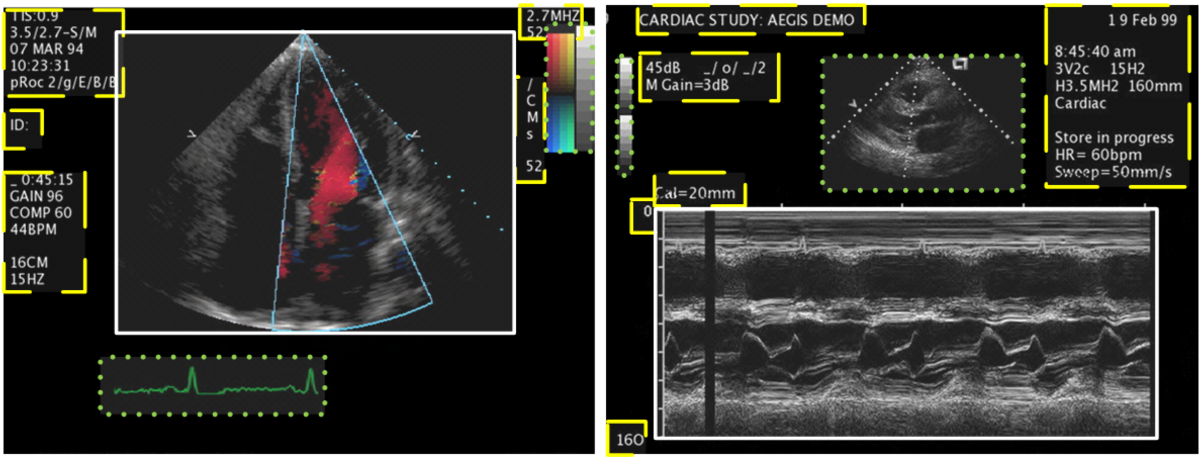


Figure 3.2: *Echocardiogram regions of a Doppler mode (on the left) and a continuous Doppler mode (on the right) captured with an Agilent and Acuson devices respectively. The white solid line contains the ultrasound, the green dotted line the auxiliary images and the yellow dashed line the text.*

Table 3.3: *Acquisition devices and cardiac affection for the real-time echocardiograms database.*

Device	Patient	Patient's Diagnosis
Sonosite SonoHeart Elite	1	Lateral myocardial infarction
	2	Medial septal myocardial infarction
	3	Ventricular septal defect
Philips ENVISOR C HD	4	Atrial septal defect
	5	Right ventricle hypertrophy and pulmonary insufficiency
	6	Mitral and tricuspid regurgitation
Philips IE33	7	Normal
	8	Tricuspid regurgitation
	9	Pulmonary insufficiency

3.3). Three sessions per device, each session corresponding with one patient, were recorded having a total of 200 videos. Each device had its specific characteristics and different qualities. First, a portable device was chosen (see SonoSite, Table 3.3) because the image quality is inferior and the manner of representation is quite different to that of non-portable devices. The other two chosen devices were of the same brand, but their quality and the form of representation were different. The image quality of both was superior to that of portable device. The selected echocardiograms were representative of typical and abnormal findings in the cardiovascular field. The physical conditions of the patients were diverse (large, medium, and slim build), as well as their ages (baby, young, middle-aged, and old).

The acquired videos had a frame rate of 25 fps and a resolution of 720x576 pixels throughout the screen. This resolution is considered high, allowing make the same diagnosis as the original image. Each frame was encoded in YUV12 format [12 bpp, 8 bpp for the luminance (Y) component and

Table 3.4: *Database for the real-time echocardiograms. Echocardiograms time and number of videos for each mode and total number of videos. B is the B mode, D the color Doppler mode, M the M mode and DP the pulsed/continuous Doppler mode.*

Patient	Total time (minutes)	% time			# videos				
		2-D	Sweep	Stop	B	D	M	DP	Total
1	28	68%	22%	10%	5	5	4	4	18
2	25	55%	27%	18%	6	6	4	4	20
3	36	64%	20%	16%	5	6	4	4	19
4	25	76%	9%	15%	8	8	5	4	25
5	30	83%	16%	1%	10	4	3	5	22
6	28	76%	10%	14%	7	9	4	5	25
7	27	52%	33%	15%	8	8	4	5	25
8	27	82%	10%	8%	7	6	24	4	41
9	30	91%	8%	1%	5	7	4	4	20

Table 3.5: *Main regions present in the dataset echocardiograms for each device brand and mode.*

Mode	2-D modes			Sweep modes		
Device	US	Text	ECG	US	Text	Auxiliary ultrasound
SonoSite	✔	✔		✔	✔	
Philips	✔	✔	✔	✔	✔	✔

4 bpp for the color components, chrominance (U and V)]. The echocardiographic sessions ranged from 25 to 36 minutes duration. The echocardiograms only contain the four basic operation modes according to the EAE described in Section 1.2. The operation modes are visualized one each time. Stops are introduced when measurements are performed or the mode is changing. The echocardiogram is composed of several fragments that correspond to a change of operation mode (videos). Table 3.4 shows for the available echocardiograms, the echocardiogram time distribution, total time and percentage of time for each mode, and the number of videos for the four modes. The modes are indicated with 2D, M, PCD and CD labels representing 2D, M, pulsed/continuous Doppler and color Doppler modes, respectively. The main regions for each device brand and visualization mode are shown in Table 3.5.

3.1.2.2 Real-Time Echocardiogram Characteristics

The echocardiogram characteristics to be taken into account are the following:

- The echocardiograms present different regions, as shown in Figure 3.3, each region having different visualization characteristics, diagnostic importance and data type.

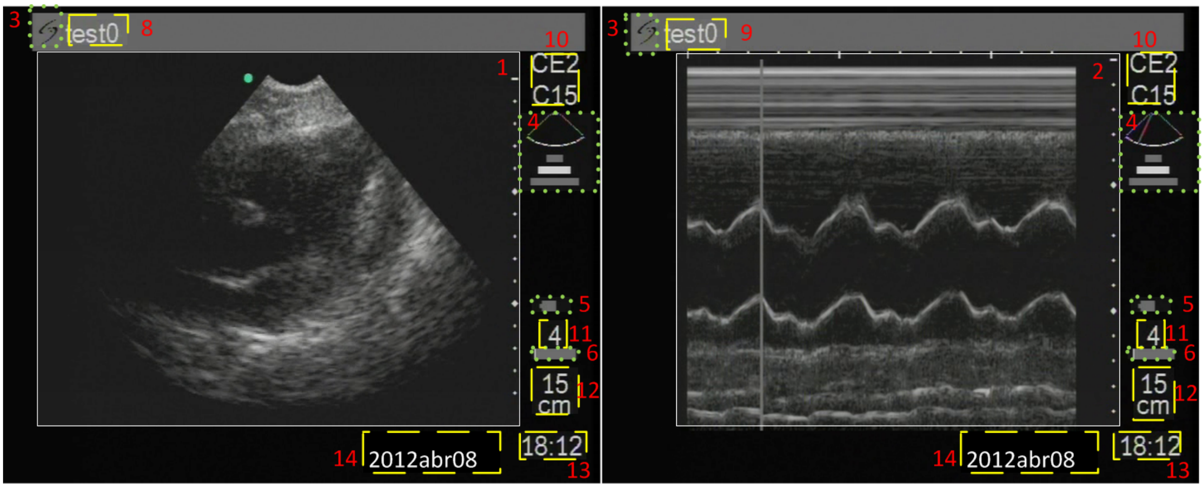
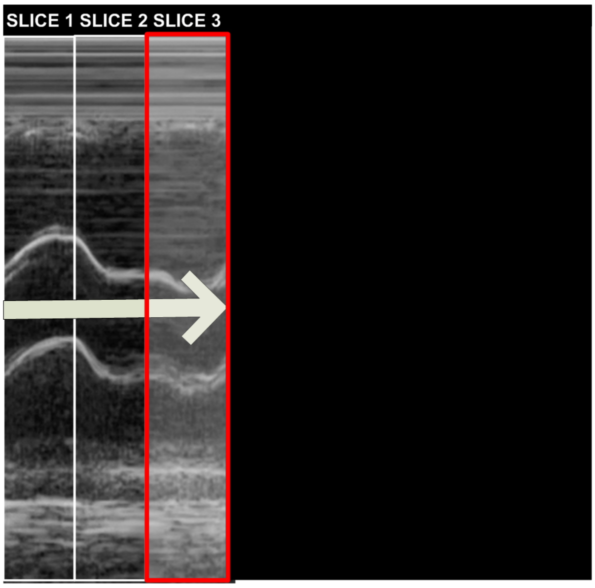
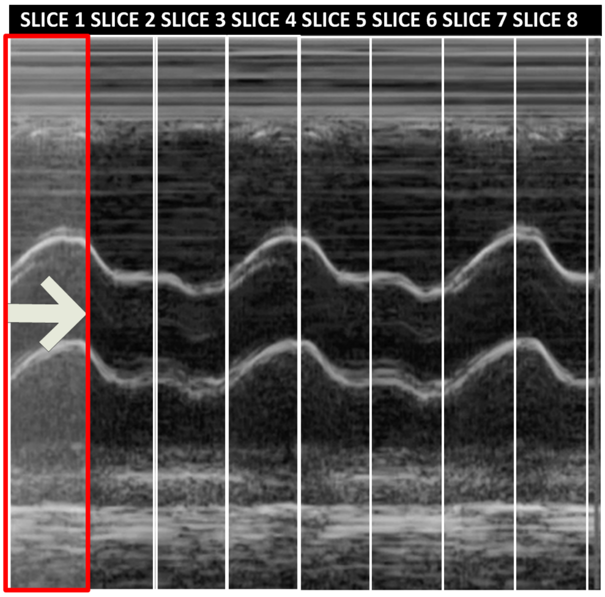


Figure 3.3: *Echocardiogram regions of a B mode (on the left) and a M mode (on the right) captured with a Sonosite device. The white solid line contains the ultrasound, the blue dotted and dashed line the ECG, the green dotted line the auxiliary images and the yellow dashed line the text. Each region is identified with a number (ID) on the figure.*

- According to their characteristics, five regions can be distinguished (see Figure 3.3):
  - Ultrasound: this is a video that contains the relevant diagnostic information. It is always present in the echocardiogram. This region changes with the mode. Furthermore, the ultrasound region can be divided into two according to its visualization characteristics:
    * 2-D ultrasound: this represents a 2-D image of the heart in movement. Examples are the B and the color Doppler mode.
    * Sweep ultrasound: this represents a temporal evolution of one cut of the heart. Examples are M and pulsed/continuous Doppler mode. The ultrasound is displayed gradually. A new slice is visualized in each frame (see Figure 3.4). When all the screen is swept, it starts again from the beginning (see Figure 3.4b). Eventually, the sweep is stopped by the cardiologist in order to take measurements, so the same image is shown in the following frames.
  - ECG: this contains the ECG signal and can be seen as a sweep video (sweep ultrasounds) or as a signal.
  - Auxiliary video: this is a video located around the ultrasound that shows an auxiliary ultrasound to help to interpret the echocardiogram.
  - Auxiliary image: this is an image located around the ultrasound that shows an auxiliary image to help to interpret the echocardiogram, such as another ultrasound image or symbols regarding the configuration. It may be merely decorative. These images are always present in all the echocardiograms, and even more than onces.
  - Sound: the echocardiogram may also incorporate sound in order to listen to the heart-beat.
  - Text: this is always present in all the echocardiograms, normally more than onces. It contains information such as patient information and date, time, configuration or measurements of the study.



(a) *Frame 3 of a sweep mode.*



(b) *Frame 9 of a sweep mode.*

Figure 3.4: *Examples of frames for a sweep mode.*

- The distribution and size of the regions change with the acquisition device and may change with the echocardiogram mode of operation. The represented or activated regions change according to the activated mode. For example, in Figure 3.3 the regions with identification numbers 1, 3, 4, 5, 6, 7, 8, 10, 11, 12, 13, 14 are activated for the B mode. For the M mode the regions 2, 3, 4, 5, 6, 9, 10, 11, 12, 13, 14 are activated. The ultrasound region is always different for each mode, thus the activated mode depends on the ultrasound region. The other regions may be common for several modes, for example regions 3, 5, 8, 10, 11, 12, 13, 14 in Figure 3.3 are in both the B and M modes.

- Depending on the data type of each region the regions that correspond with the activated mode of operation may change with time or may hardly ever change. The auxiliary image and text remain invariant and only change occasionally. However, the rest of the regions change each few seconds and thus synchronism between them is necessary.

- Some regions contain color information relevant for the diagnosis while others do not. For example, the color Doppler modes have relevant information in color in the ultrasound region and in the color scale region.

- The echocardiogram can be seen like a stored video. It can be stopped and run forward and back. This only affects the regions that change with time.

In Table 3.6 is summarized the regions that may be presented in an echocardiogram, its data type and if the regions need synchronism between them.

Table 3.6: *Data type of the echocardiogram regions for real-time transmission.*

Regions	Data type	Synchronism
Ultrasound	Video or sweep video	Yes
ECG	Signal or sweep video	Yes
Auxiliary video	Video	Yes
Auxiliary image	Image	No
Sound	Audio	Yes
Text	Text	No

3.2 Clinical Evaluation Methodology for Compression Recommendations

This section presents an accurate clinical evaluation methodology for compressed echocardiograms and the four basic modes of operation according to the EAE recommendations. A two phase evaluation is proposed in order to simplify the evaluation process and save time compared to other accurate methodologies presented in the literature. The first phase is a semi-blind test to discriminate two transmission rates. The second phase is a more extensive blind test for the two transmission rates that are selected in the previous test. The evaluation leads to a recommendation for the compression of echocardiograms. Both tests are described in the following sections.

3.2.1 First Phase: Semi-blind Test

The objective of this test is to determine in a fast and subjective way whether the cardiologist would be able or not to make the same diagnosis with the original and the compressed video. This test consists of three different parts, see Figure 3.5. The first part permits us to measure the cardiologist's opinion about the similarity between the compressed echocardiogram and the original video. The second part is a question about whether the cardiologist's diagnosis would be the same for both videos, the original and the compressed. The third part invites comments, if any. These questions are answered for each echocardiogram mode.

1. Measure of similarity between the original X mode video and the compressed one.

1: very different	2: different	3: acceptable	4: similar	5: identical

2. Would you give the same diagnosis with both videos for the X mode?

YES	NO

3. Comments.

Figure 3.5: *Semi-blind test: comparison of compressed echocardiogram with the original video.*

In order to have an estimation that directly reflects the distortion in the diagnosis content of every mode in the echocardiogram, a clinical distortion index for the semi-blind test (CDI_{SB}) is calculated with all the information collected for each mode, cardiologist and transmission rate. This is defined as

$$CDI_{SB} = \frac{1}{2} \cdot \frac{(5-C)}{5} + \frac{1}{2} \cdot (1-D) \tag{3.1}$$

where C is the measurement of similarity between the original and the compressed video (1-5) and D is the answer to the boolean question about the diagnosis (1-YES, 0-NO) for the mode being evaluated (see Figure 3.5).

The CDI_{SB} is directly related to the clinical distortion of the echocardiogram. The lower the value of CDI_{SB}, the lower the distortion in the diagnosis content of the compressed echocardiogram. The CDI_{SB} could be grabbed into quality ranges thus making it possible to classify the echocardiograms. Cardiologists involved in the study considered that it would be more practical to split the CDI_{SB} values (0-0.9) into three ranges. The three defined ranges are:

- The same diagnosis is possible ($D = 1$) with acceptable quality ($C \geq 3$): $CDI_{SB} \leq 0.2$
- The same diagnosis is possible ($D = 1$) but with low quality ($C < 3$): $0.2 < CDI_{SB} \leq 0.4$
- The same diagnosis is not possible ($D = 0$) with low quality ($C < 3$): $CDI_{SB} > 0.7$

It is important to note that there are no CDI_{SB} values between 0.4 and 0.7. It is because the cardiologists decided that it is not possible to have that the same diagnosis is not possible ($D = 0$) with acceptable quality ($C \geq 3$). This implies that if the same diagnosis is not possible the quality has to be unacceptable. The CDI_{SB} values are used to discard the transmission rates that are in the not acceptable range ($CDI_{SB} > 0.4$) and to obtain the two transmission rates that are between acceptable or low quality but with the same diagnosis ($CDI_{SB} \leq 0.2$ and 0.2$\leq CDI_{SB} \leq 0.4$) for each mode. In order to assess the test at least three cardiologist should participate and for each mode several videos of different patients and ultrasound devices should be evaluated. Once the CDI_{SB} has been calculated for all the patients, the two transmission rates are selected as follows: the upper rate is the first transmission rate with all the CDI_{SB} values of the different patients equal or lower than 0.2 and the lower rate is the immediately inferior. These two rates are the ones selected for a deep evaluation using a more extensive test, the blind test.

3.2.2 Second Phase: Blind Test

This test corresponds to the same blind test proposed in [56] and consists of three different parts, which are much more extensive than in the semi-blind test. We have maintained this test because it is very complete, and now with the advantage that we just have to evaluate the transmission rates selected in the previous test, saving in this way a lot of time in the evaluation process. The objective of this test is evaluate whether the same diagnostic is possible or not with both videos, the original and the compressed video. In Figures 3.6, 3.7 and 3.8 the three parts of the test for the four basic modes of operation are shown. The first part, Figure 3.6, shows the overall score given by cardiologists to the general video quality. The second part, Figures 3.7 and 3.8, consists of some questions that cardiologists have to complete based on the interpretations of the structures and flows measured in a standard echocardiogram examination (the modes for each interpretation are shown in brackets). If these interpretations are the same as for the original video the same diagnostic is possible. Finally, there is a part for comments from the specialists, if any.

1a. General quality score for the B mode.

1: very bad	2: bad	3: tolerable	4: good	5: excellent

1b. General quality score for the M mode.

1: very bad	2: bad	3: tolerable	4: good	5: excellent

1c. General quality score for the color Doppler mode.

1: very bad	2: bad	3: tolerable	4: good	5: excellent

1d. General quality score for the pulse and continuous Doppler mode.

1: very bad	2: bad	3: tolerable	4: good	5: excellent

Figure 3.6: *Blind test: part I.*

The clinical distortion index for the blind test (CDI_B) is defined as

$$CDI_B = max\left\{\frac{Q_o - Q_r}{2 \cdot Q_o}, 0\right\} + \frac{\sum |sgn\,(I_{0i} - I_{ri})|}{2 \cdot K} \tag{3.2}$$

where Q_o and Q_r are the general quality scores (see Figure 3.6) of the original and compressed videos for the mode being evaluated, respectively. For the scores rating special attention has been paid to the sharpness and definition of the edges of structures, their clear separation from other structures and the clarity in blood floods. I_{oi} and I_{ri} are the interpretation of the ith parameter of the original and reconstructed videos, respectively. These interpretations are translated into numeric factors in order to be used in the equation. A zero value is assigned if the cardiologist can not answer the question due to bad video quality. The rest of the questions are assigned a value larger than 0. K is the number of questions that contribute to each echocardiographic mode. K is nine for the B mode, six for the M mode, five for the color Doppler mode and six for the pulsed/continuous Doppler mode.

As described before at least three cardiologists should participate and for each mode several videos of different patients and ultrasound devices should be evaluated. As was established in [56], the maximum CDI value obtained by a compressed echocardiogram in order to be considered acceptable should be at most 0.25, being the CDI the mean between CDI_B and CDI_{SB}. In other words, the minimum transmission rate to guarantee the clinical quality will be the lower with all the CDI values lower than 0.25. For example, let consider we evaluate a color Doppler mode

2. Provided with an interpretation.

<table>
<tr><td rowspan="40">B and M study</td><td colspan="2" rowspan="3">(B, M) Aortic root (transversal diameter)</td><td colspan="3">Normal</td></tr>
<tr><td colspan="3">Dilated</td></tr>
<tr><td colspan="3">Dissected</td></tr>
<tr><td colspan="2" rowspan="3">(B, M) Left atrium (transversal diameter)</td><td colspan="3">Normal</td></tr>
<tr><td colspan="3">Dilated</td></tr>
<tr><td colspan="3">Intra-arterial mass-clot</td></tr>
<tr><td rowspan="4">Left ventricle</td><td rowspan="2">LVDD (M)</td><td colspan="3">Normal</td></tr>
<tr><td colspan="3">Dilated</td></tr>
<tr><td rowspan="2">LVSD (M)</td><td colspan="3">Normal</td></tr>
<tr><td colspan="3">Dilated</td></tr>
<tr><td colspan="2" rowspan="6">(B) Global contractility</td><td colspan="3">Hyperkinetic</td></tr>
<tr><td colspan="3">Normal</td></tr>
<tr><td rowspan="3">Depressed</td><td colspan="2">Light</td></tr>
<tr><td colspan="2">Moderate</td></tr>
<tr><td colspan="2">Severe</td></tr>
<tr><td colspan="3">Asynergy</td></tr>
<tr><td rowspan="16">Left ventricle wall thickness</td><td rowspan="8">(M) IVS</td><td colspan="3">Thin</td></tr>
<tr><td colspan="3">Normal</td></tr>
<tr><td rowspan="6">Left ventricular hypertrophy</td><td rowspan="3">C_a</td><td>Light</td></tr>
<tr><td>Moderate</td></tr>
<tr><td>Severe</td></tr>
<tr><td rowspan="3">A^b</td><td>Light</td></tr>
<tr><td>Moderate</td></tr>
<tr><td>Severe</td></tr>
<tr><td rowspan="8">(M) PW</td><td colspan="3">Thin</td></tr>
<tr><td colspan="3">Normal</td></tr>
<tr><td rowspan="6">Left ventricular hypertrophy</td><td rowspan="3">C_a</td><td>Light</td></tr>
<tr><td>Moderate</td></tr>
<tr><td>Severe</td></tr>
<tr><td rowspan="3">A^b</td><td>Light</td></tr>
<tr><td>Moderate</td></tr>
<tr><td>Severe</td></tr>
<tr><td rowspan="6">Valve Morphology</td><td rowspan="2">Mitral (B)</td><td colspan="3">Normal</td></tr>
<tr><td colspan="3">Abnormal</td></tr>
<tr><td rowspan="2">Aortic (B)</td><td colspan="3">Normal</td></tr>
<tr><td colspan="3">Abnormal</td></tr>
<tr><td rowspan="2">Tricuspid (B)</td><td colspan="3">Normal</td></tr>
<tr><td colspan="3">Abnormal</td></tr>
<tr><td colspan="2" rowspan="2">Pericardial effusion (B)</td><td colspan="3">Yes</td></tr>
<tr><td colspan="3">No</td></tr>
</table>

C_a: Concentric, A^b: Asymetric

Figure 3.7: *Blind test: part II.*

3. Provided with an interpretation.

<table>
<tr><td rowspan="31">Doppler Study</td><td rowspan="7">(PCD) Pulmonary flow</td><td rowspan="7" colspan="2">Max. velocity</td><td colspan="2">Normal</td></tr>
<tr><td rowspan="3">Sten.</td><td>Light</td></tr>
<tr><td>Moderate</td></tr>
<tr><td>Severe</td></tr>
<tr><td rowspan="3">Ins.</td><td>Light</td></tr>
<tr><td>Moderate</td></tr>
<tr><td>Severe</td></tr>
<tr><td rowspan="8">Mitral Flow</td><td rowspan="4">Syst.</td><td rowspan="4">Regurgitation (PCD, CD)</td><td rowspan="3">Yes</td><td>Light</td></tr>
<tr><td>Moderate</td></tr>
<tr><td>Severa</td></tr>
<tr><td colspan="2">No</td></tr>
<tr><td rowspan="4">Diast.</td><td rowspan="4">E. wave, A. pattern (PCD)</td><td colspan="2">Normal</td></tr>
<tr><td colspan="2">Pseudonormal</td></tr>
<tr><td colspan="2">Relaxation Abnormality</td></tr>
<tr><td colspan="2">Restrictive</td></tr>
<tr><td rowspan="8">Aortic flow</td><td rowspan="4">Syst.</td><td rowspan="4">Max. Velocity (PCD)</td><td colspan="2">Normal</td></tr>
<tr><td rowspan="3">Sten.</td><td>Light</td></tr>
<tr><td>Moderate</td></tr>
<tr><td>Severa</td></tr>
<tr><td rowspan="4">Diast.</td><td rowspan="4">Regurgitation (PCD, CD)</td><td rowspan="3">Yes</td><td>Light</td></tr>
<tr><td>Moderate</td></tr>
<tr><td>Severe</td></tr>
<tr><td colspan="2">No</td></tr>
<tr><td rowspan="4">Tricuspid flow</td><td rowspan="4" colspan="2">Regurgitation (PCD, CD)</td><td rowspan="3">Yes</td><td>Light</td></tr>
<tr><td>Moderate</td></tr>
<tr><td>Severe</td></tr>
<tr><td colspan="2">No</td></tr>
<tr><td rowspan="4">Septal defects</td><td rowspan="2" colspan="2">ASD (2D,CD)</td><td colspan="2">Yes</td></tr>
<tr><td colspan="2">No</td></tr>
<tr><td rowspan="2" colspan="2">VSD (2D,CD)</td><td colspan="2">Yes</td></tr>
<tr><td colspan="2">No</td></tr>
</table>

Figure 3.8: *Blind test: part III.*

($K = 5$), in which the original echocardiogram has general quality of 4 ($Q_o = 4$), the compressed echocardiogram of 3 ($Q_r = 3$), and for the first four interpretations the values are the same for both echocardiograms and for the last interpretation the original value is septal defect yes ($I_{o4} = 2$) and with the compressed echocardiogram the cardiologist can not answer ($I_{r4} = 0$). The value of CDI_B would be 0.325 [(4-3)/(2·4) + (0+0+0+0+2)/(2·5)]. If the CDI_{SB} value is lower than 0.125, the CDI value would be lower than 0.25 thereby acceptable, other case the CDI value would be not acceptable.

3.3 Compression for Storage

An image format for storage purpose is proposed in this section. As mentioned in Chapter 2, the proposed storage file has to be interoperable with DICOM and has to take into account the DICOM file format. The final file follows the DICOM format, and consequently has two parts (see Section 2.1.1): the DICOM header and the proposed image format. No changes have to be made in the DICOM header to integrate the proposed image format in the DICOM standard. It is only necessary to define a "Transfer Syntax Unique Identifier" for the proposed format in the standard in order to identify that the image format used is the proposed format.

The design of the proposed image format takes into account the characteristics of the stored echocardiograms described previously and the encoding recommendations in Chapter 2. Furthermore, we take advantage of the facilities of segmentation already incorporated in the acquisition devices. The acquisition device provides the regions and their type: ultrasound, text or auxiliary.

An appropriate selection of the compression algorithms for each region is very important in order to save storage space. In a previous study we proposed SPIHT algorithms [118]. However, although good results were obtained, these algorithms were rejected because they are not included in the DICOM standard. JPEG 2000 was chosen because similar results to those of the SPIHT algorithm were achieved and it is included in the DICOM standard.

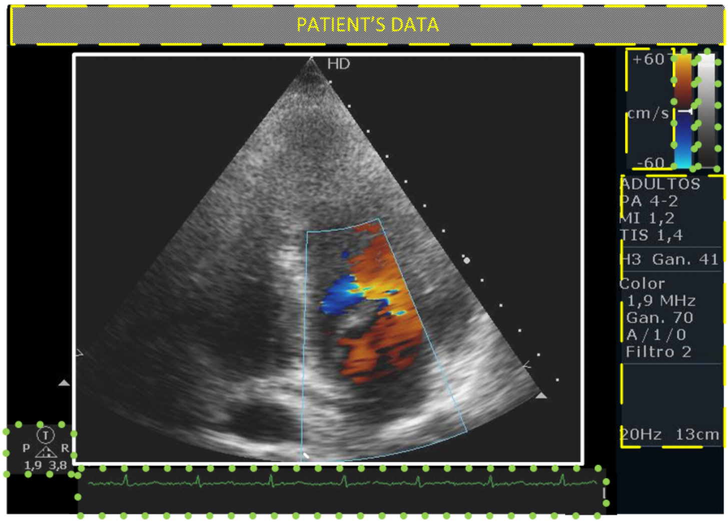
Since each echocardiogram has a different distribution of the regions, it is necessary to define the configuration for each stored echocardiogram. A ROI coding as the Maxshift method specified in JPEG 2000 [35] or the context based coding [119] are not proposed because they do not support text compression. It is important to compress the text without losing quality and efficiently since it can contain relevant information for the diagnosis. Furthermore, if the regions are encoded separately, the image can be easily edited after being stored. The edition of the image is very useful for the laboratories, because important information for the diagnosis can be added afterwards. We proposed a first approach in [118], where some headers were added to the global file and the encoded regions to indicate global configuration, position and size. This solution is highly efficient from the point of view of saving space. However, the solution is neither flexible nor portable. Extensible Markup Language (XML) [120] is a markup language that defines a set of rules for encoding information. It uses human language, not computer language, which is readable and understandable. It is an expandable language, meaning that new tags can be created or previously created tags can be used. It is also extremely portable. Furthermore, XML is similar to the DICOM headers system [121]. For these reasons, XML is proposed for the configuration of the image format.

The final proposed image format has two parts: an XML file, where the information relating to the configuration of the regions is included, and the encoded regions. Both parts are described below.

```
<?xml version="1.0" encoding="UTF-8"?>
<!ELEMENT format (tsize?,region+)>

<!ELEMENT tsize EMPTY>
<!ATTLIST tsize w CDATA #REQUIRED>
<!ATTLIST tsize h CDATA #REQUIRED>

<!ELEMENT region (pos,(roi|img|text))>

<!ELEMENT pos EMPTY>
<!ATTLIST pos x0 CDATA #REQUIRED>
<!ATTLIST pos y0 CDATA #REQUIRED>
<!ATTLIST pos x1 CDATA #IMPLIED>
<!ATTLIST pos y1 CDATA #IMPLIED>

<!ELEMENT roi (size+)>
<!ELEMENT size (#PCDATA)>
<!ATTLIST roi cod CDATA "jpeg2000">
<!ATTLIST roi frames CDATA #IMPLIED>

<!ELEMENT img EMPTY>
<!ATTLIST img L  CDATA #REQUIRED>
<!ATTLIST img cod CDATA "jpeg2000">

<!ELEMENT text (#PCDATA)>
```

Figure 3.9: *DTD file for the configuration of the stored echocardiograms.*

- **XML file:** The XML syntax has to be defined, for example, with a Document Type Definition (DTD) or an XML Schema file [122]. The proposed DTD file is defined in Figure 3.9. An example of an XML file for three regions of the Doppler mode is shown in Figure 3.10. The XML file contains the following information, including in brackets the XML fields:
  - The size of the whole image (*tsize*: w, h), although this information is already in the DICOM header.
  - The regions' configuration (*region*), one for each region that appears in the image. The type of regions are: ultrasound or ROI (*roi*), auxiliary image (*img*) and text (*text*). All the regions have a rectangular shape.
  - The position of the regions (*pos*). The initial position ($x0$, $y0$) has to be defined for all the types of region and also the final position ($x1$, $y1$) except for the text region since its size is adjusted to the available space.
  - The type of codification (*cod*) may be indicated for the ultrasound and image regions, the default codec being JPEG 2000.
  - The sizes in bytes of the ultrasound (*size*) and image (L) regions have to be specified in bytes.
  - In the case of having several frames, the ultrasound is the only region that changes in every frame. The ultrasound regions of every frame share the same configuration, so



(a) *Echocardiogram image of a color Doppler mode.*

```xml
<?xml version="1.0" encoding="UTF-8"?>
<!DOCTYPE configuration SYSTEM "roiformat.dtd">
<format>
        <region>
                <pos x0="44" y0="75" x1="564" y1="597"></pos>
                <roi><size>33893</size></roi>
        </region>
        <region>
                <pos x0="0" y0="0"></pos>
                <text>PATIENT'S DATA</text>
        </region>
        <region>
                <pos x0="77" y0="574" x1="615" y1="573"></pos>
                <img L="149"></img>
        </region>
</format>
```

(b) *XML.*

Figure 3.10: *XML example for the Philips device in the image.*

only one region (roi) has to be indicated in the XML file. The number of frames is indicated in the ultrasound region ($frames$). The field size ($size$) of every ultrasound frame can be added in the case that the size changes for each one, otherwise it only has to appear once.

– The text regions ($text$) include the text in the XML file.

Although some of the XML information may already be included in the DICOM header, it is necessary to include it in the image format to have an image with a self-contained format. This means that the image can be decoded without any additional data. If the acquisition devices incorporate the proposed format they have to provide the XML file. This task is not difficult since there are similar headers in the DICOM file that the devices already provide, such as the file size or the calibration regions.

- **Encoded regions:** The second part is the encoded regions, except the text that is included in the XML file. The regions appear in the same order as in the XML file. The auxiliary images and ultrasound image or images, if there are several frames, are compressed with JPEG 2000 by default, which is included in the DICOM standard. Other compression image formats can be used. The method is indicated in the configuration part. As a result of the compression method, each image region may be compressed with different quality according to its diagnostic relevance. This encoding process does not add complexity to the acquisition device since the codification method is already included in devices complying with the DICOM standard.

3.4 Results and Discussion for Compression for Storage

A clinical evaluation of the proposed compression is not necessary since previous evaluations have been carried out for ultrasound images compressed with JPEG 2000 [44]. Thus, in this section we focus on the compression rates obtained with the storage format proposed in this Thesis and the storage gain compared to compression without distinguishing regions. An application has been developed in Matlab that converts the typical medical images available in laboratories and hospitals into the proposed storage format and vice versa. This functionality allows studies to be stored with the proposed format, thus saving disk space and reducing the transmission time of the studies. This tool can also be used for the evaluation of the results in this section. Furthermore, the tool allows the image that has been stored in the proposed format or other formats to be displayed and edited, and for measurements to be performed. The tool takes advantages of the proposed format based on regions.

3.4.1 Stored Echocardiogram Tool: Format Converter, Display and Measure

The main functionality of the tool is the conversion into the proposed format (see Figure 3.11). The application input is a DICOM file or another image file format used for medical images [29]. It is important that the input image has sufficient quality to be able to extract the text and for the doctor to make a diagnosis. If the text is highly distorted it is difficult to recognize it. The transformation to the proposed format has three main parts that are represented in Figure 3.11 and described below.

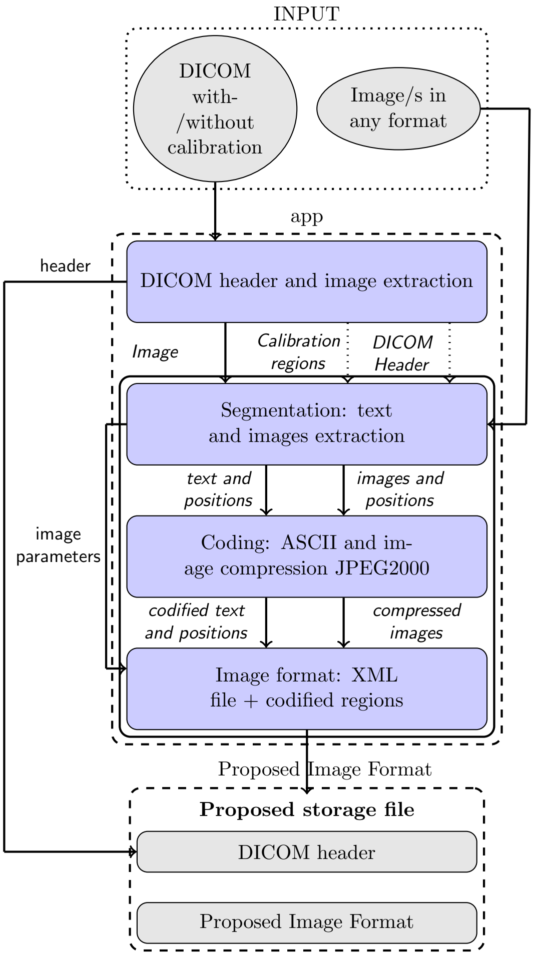


Figure 3.11: *Process of converting the stored image into the proposed storage file.*

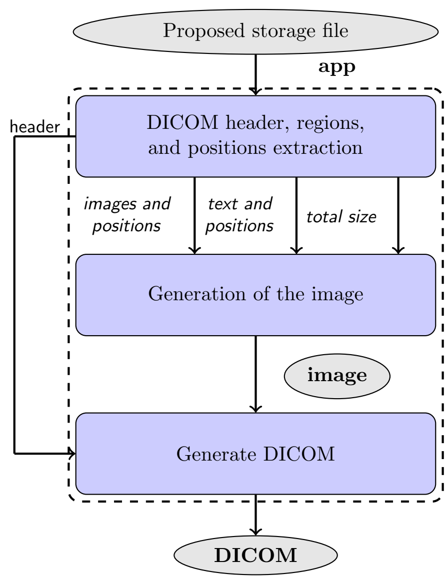


Figure 3.12: *Process of converting the proposed format into another format.*

- Header and image extraction. In the case of having a DICOM file, it is necessary to extract the header for the final format and the image or images for the following steps. In many cases, depending on the device, the DICOM file incorporates a calibration section which indicates where the ultrasound region is, and consequently the segmentation process is simplified. The header can also be useful because it contains information about the device.

- Proposed image format. This part has three steps: the segmentation of the images, the coding of the regions and the composition of the image format (XML file and the coded regions). The segmentation identifies the regions and the type, ultrasound, text or auxiliary, of the input images. In the case of having several frames, as the only part that changes is the ultrasound region, it is necessary to extract the text and the auxiliary images of only one frame, and extract the ultrasound region of all the frames. Furthermore, it is also necessary to know if the images (ultrasound and auxiliary regions) contain color or not and to identify the text to convert it into American Standard Code for Information Interchange (ASCII) code. In the case of having the calibration regions, the process is simplified as it is known where the ultrasound and ECG regions are. It is also possible, knowing the device, to know where the regions are, because for the same device and mode the regions are located in the same position. Our tool is able to divide the studies for all the echocardiograms included in the available database with a standard configuration file and without identifying the device. The text is recognized by Optical Character Recognition (OCR), using an open-source software GOCR [123]. The tool also allows the doctor to select the quality of the regions having the possibility of changing the quality if it is not adequate for the diagnosis. A standard quality has been set as default value.

- The last part is to generate the final file with the proposed storage file (DICOM header and the proposed image format).

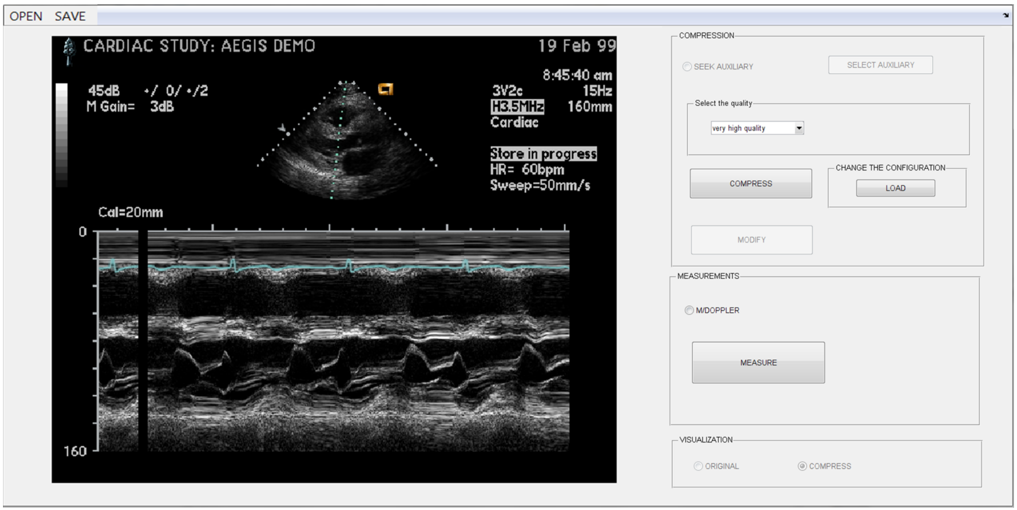
The another main functionality of the application is the reverse process. This permits converting studies stored in the proposed format into one or several images and into a DICOM file with other image formats. This functionality is useful to display the images or if it is necessary to interchange the studies with other systems not compatible with the proposed format. The steps for the transformation process are shown in Figure 3.12 and described below:

- Extraction of the DICOM header, regions and positions. The first step is to extract the images and their position, the texts and their position, the total size, and the number of frames from the proposed format (XML file and encoded regions).

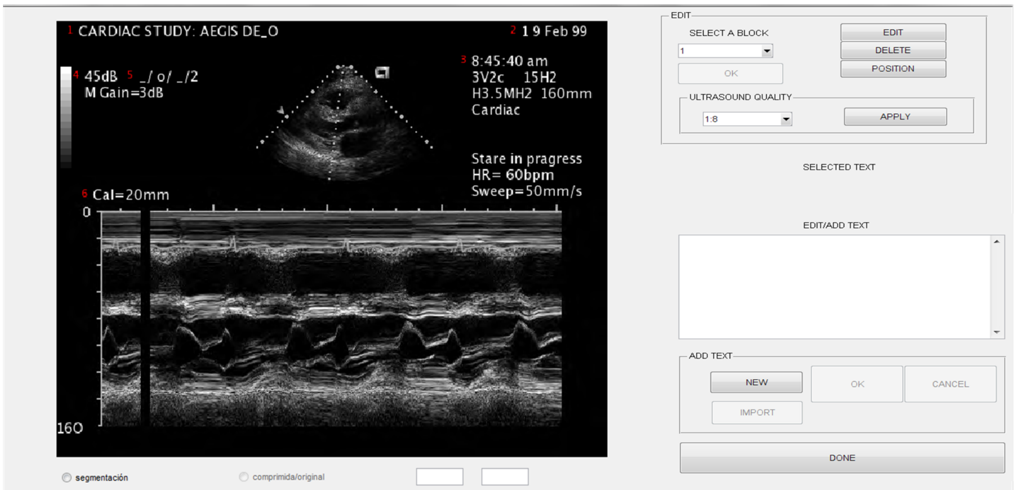
- Generation of the image. The image is generated with the specified format and with all the information provided by the previous step.

- Generation of the DICOM file. The image is joined with the header extracted in the first step and the DICOM file is generated.

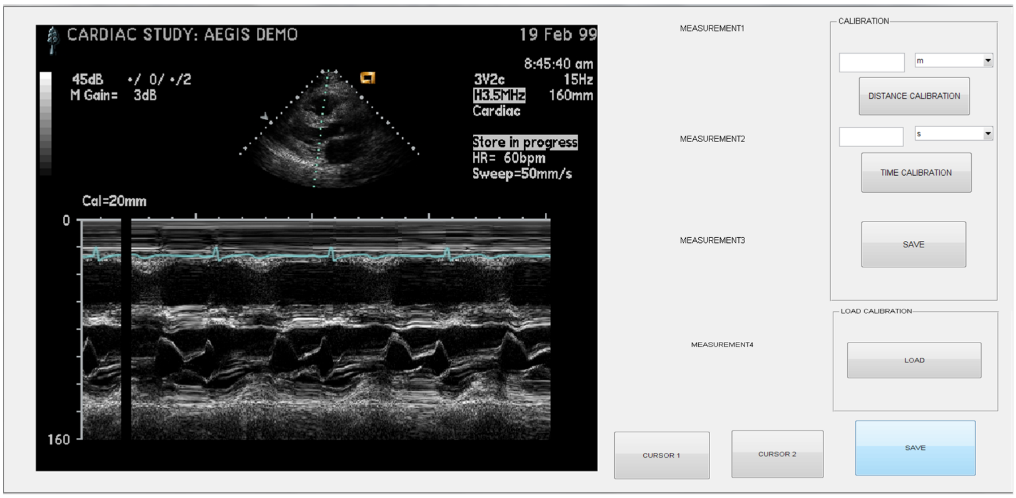
Moreover, more functionalities have been added to achieve a more complete application that fulfills the requirements of laboratories and takes advantages of the proposed format based on regions. The application has been designed with the collaboration of three expert cardiologists.



(a) *Main window of the tool.*



(b) *Edition window the tool.*



(c) *Measurements window of the tool.*

Figure 3.13: *Screen shots of the stored echocardiogram tool: compression, edition and measurements.*

Screen shots of the application are shown in Figure 3.13. The additional implemented functionalities are the following:

- Visualization of the studies (Figure 3.13a). The tool is compatible with DICOM, images in other formats [29] and the proposed storage format.
- A tool to modify, add and remove regions (Figure 3.13b). The tool is also able to change the text or add more regions with text. This functionality is very helpful since it allows the specialist doctor to add or delete information and consequently store more complete studies.
- Measurements tool (Figure 3.13c). This is a very useful functionality to measure after the acquisition. The tool is able to charge the calibration parameters, if they are available in the DICOM header, or to perform the calibration and store these parameters in the DICOM header. It is also able to store the study including the measurements.

3.4.2 Results for Compression of Stored Echocardiograms

The compression gain can be calculated previously to the compression process. The compression rate (CR) for the proposed compression format with respect to the raw image with and without including the DICOM header follows the expression:

$$CR = \frac{width_T * height_T * N * 3 * 8\,(+\,header_{DICOM})}{header + width_U * height_U * N * bpp\,(+\,header_{DICOM})} \tag{3.3}$$

$width_T$ and $height_T$ are the total size of the image, N is the number of frames, 3 corresponds to the three color components, and 8 to the bits per pixels. The *header* is the header size of the proposed format including the XML file and auxiliary images, $width_U$ and $height_U$ are the size of the ultrasound region, and *bpp* are the bits per pixels or bit rate for the compressed ultrasound region. The $header_{DICOM}$ is the DICOM header that may or may not be considered. The DICOM header is taken into account if the compression rate of the whole file needs to be calculated. However, the DICOM header is not included in the expression if the compression rate of the image format is to be evaluated.

The following expression 3.4 is the percentage of gain in storage space of the proposed compression with respect to the compression study using JPEG 2000, which does not distinguish regions. The DICOM header may or may not be included. The ultrasound regions are compressed with the same quality for both methods.

$$Gain = \frac{width_T * height_T * N * bpp - header - width_U * height_U * N * bpp}{header + width_U * height_U * N * bpp\,(+\,header_{DICOM})} * 100 \tag{3.4}$$

where the parameters are the same as those defined for CR.

In order to evaluate the improvement of the proposed compression format compared to JPEG 2000, the echocardiograms database has been compressed with both methods. Table 3.7 shows for each echocardiogram device and mode available, the ultrasound size, average header size (XML file and auxiliary images) for one frame, the CR for the proposed format and the percentage of gain (G)

Table 3.7: *Parameters and compression results for the echocardiogram devices and modes of the database.*

Device	Mode	Header (Bytes)	ROI Size	CR/CR_{DICOM}		$G/G_{DICOM}(\%)$	
				1 frame	16 frame	1 frame	16 frame
Agilent/ Acuson	M	4401	576x295	31/19	-	28/16	-
	DP	5036	569x295	30/20	-	24/15	-
	D	2746	404x417	29/21	32/31	20/14	32/32
	B	825	551x480	27/20	28/27	14/9	16/16
Philips	M	6161	692x387	35/25	-	48/32	-
	DP	6838	641x369	39/25	-	61/37	-
	D	3510	538x521	36/21	39/37	49/27	62/60
	B	2729	442x521	39/27	42/41	62/41	75/73
Siemens	M	11489	1024x504	34/28	-	41/32	-
	DP	11637	1024x499	33/26	-	35/28	-
	D	3227	845x739	28/22	29/29	19/14	23/23
	B	3481	824x739	29/23	31/30	22/17	27/27

of the proposed format with respect to JPEG 2000. The last two values have been calculated for the average of all the frames stored separately and considering 16 frames of the same visualization stored in the same file and sharing the configuration file. Furthermore, the results not taking into account the DICOM header and those taking into account the DICOM header are shown. The ultrasound regions have been compressed with 1 bpp (bit per pixel) for both formats and the auxiliary regions with 0.5 bpp for the proposed format. 1 bpp and 0.5 bpp have been selected because these bit rates have shown good clinical quality for ultrasound images compressed with JPEG 2000 [44].

In order to provide a visual inspection of the proposed method, Figure 3.14 shows the original images and the images after segmentation, compression and decompression. The ultrasound region has been compressed with 1 bpp and the auxiliary regions with 0.5 bpp for both images. The top images were acquired with an Agilent device corresponding to the Color Doppler mode. The PSNR of the ultrasound part is about 40 dB. The size for the proposed format is 24 kbytes and the original size is 774 kbytes, having a CR of 32. The bottom images were acquired with Acuson device corresponding to the M mode. The PSNR of the ultrasound region is about 35 dB. The size of the proposed format is 26 kbytes instead of 788 kbytes, having a CR of 30.

3.4.3 Discussion for Compression of Stored Echocardiograms

According to the compression rate and gain expressions given in the previous section, the compression efficiency mainly depends on the header length (XML file and auxiliary images) and ultrasound size (ROI), and consequently depends on the acquisition devices. For example, in Table 3.7, the Agilent/Acuson device and B mode shows the worst compression results (CR = 27 for 1 frame). This is due to the ROI size (551x480 of 600x430). The ROI covers almost the whole image. On the other hand, the best compression results are for the Philips device and B mode (CR =

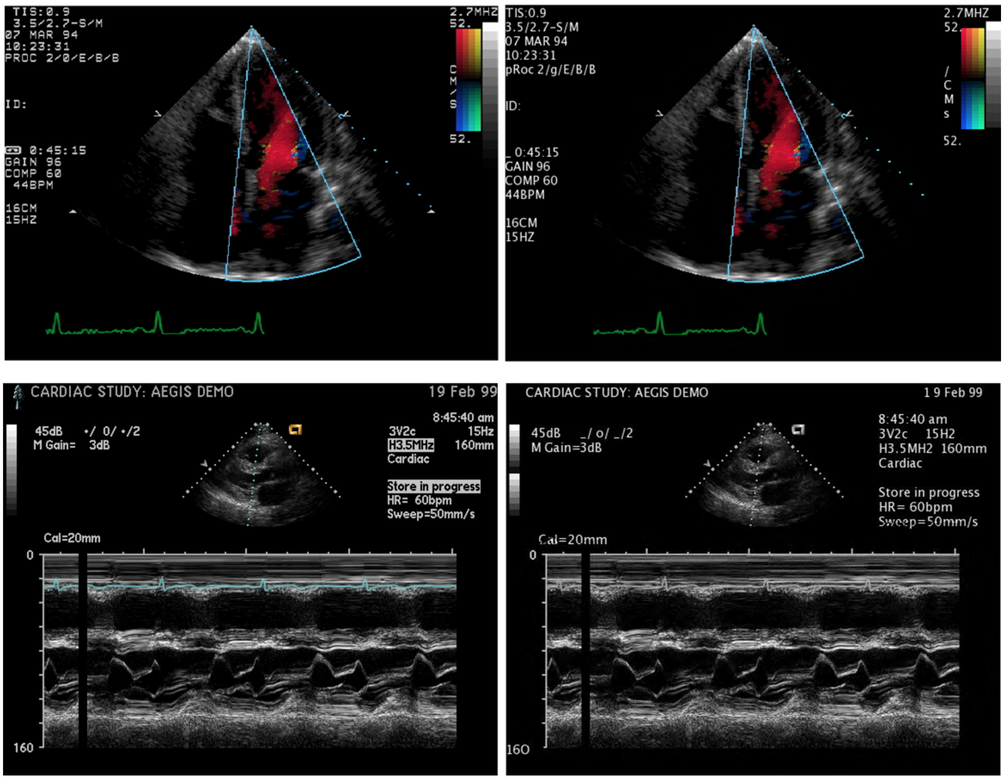


Figure 3.14: *Samples of echocardiogram images before and after compression with the proposed method. Images on the left are the original images and on the right the reconstructed images with the proposed method. The upper images were acquired with an Agilent device and the lower images with an Acuson device.*

39 for 1 frame) due to its small ROI size (442x521 of 800x564) and header length (2729 bytes). Consequently, the smaller the ROI and header, the better the compression results.

Another important factor is the number of frames (N). If more frames are compressed together, sharing the same headers, better compression efficiency is achieved. This is easily visualized if we compare the results taking into account and without taking into account the DICOM header (see Table 3.7). The compression results considering the DICOM header are worse than those not considering it. However, the compression results depend more on the DICOM header when only one frame is stored in a DICOM file.

In conclusion, we can observe that for a typical quality the proposed format overcomes the JPEG 2000 standard for all the echocardiograms of the database. Furthermore, the obtained CR enhances the results of the ROI coding included in the JPEG 2000 standard, Maxshift method, that obtains a CR of 16 [35] for medical images with smallest ROI region than the ultrasound. Furthermore, although the compression results depend on the acquisition device and mode, a saving in the storage space is achieved for all the available devices and modes, which are representative of typical echocardiogram image distribution. It is expected that similar results will be obtained for other image codecs.

The proposed format is specially designed for ultrasound images, but can also be applied to other medical image modalities. However, the compression performance depends on the image visualization. Thus, the compression results will have to be evaluated for each image modality.

3.5 Compression for Real-time Transmission

An encoded algorithm is presented for each data type according to the previously presented echocardiogram characteristics. The encoding recommendations in Chapter 2 are taken into account. As the ultrasound region is the largest and the most important for the diagnosis, it is necessary to take special care in the design of the compression method for this region.

In a previous study [124] we proposed a compression method for the ultrasound region based on visualization modes (sweep and 2-D modes) and the use of SPIHT algorithms. The proposed approach was compared with the typically used video approaches, Xvid and H.264/AVC, showing better results for the compression by visualization modes in terms of PSNR for all the echocardiogram modes. For this reason, the visualization characteristics have been taken into account and SPIHT algorithms are proposed in this thesis for compression for real-time transmission. Some changes have been added to the proposal in [124] for the 2-D modes since the database was increased and the design was optimized for a wide set of echocardiograms. In [124], Run Length Encoding (RLE) was proposed for the compression of the color components and the 2-D modes. This algorithm is efficient for echocardiograms acquired with low color resolution devices such as the Sonosite device used in [124]. Nevertheless, it is not efficient for devices with high color resolution.

The final encoded recommendations for every data type is summarize in Table 3.8 and described bellow.

- **Video**. For the video compression, 3-D SPIHT [62] is proposed due to its good performance. The resolution time, time between frames encoded together, is the corresponding to 16 frames.

As the available database has 25 frames per second, the resolution time for the database is of 0.64 seconds.

- **Sweep mode**. The visualization characteristics of the sweep modes have been previously described. The image is visualized gradually. A new slice appears in each frame, remaining the rest invariant. The proposed approach compresses the new slice each frame. The slices are compressed with 2-D SPIHT due to the reasons named in Chapter 2. The images can contain color or not. If the image contains color 2-D CSPIHT [69] is proposed. On the contrary, if the image does not contain color, 2-D SPIHT [41] is proposed. The compression efficiency can be affected by the slice width. The slice height is large enough. A minimum slice width of 32 pixels has been considered suitable. If the slice width is less than 32 pixels, several slices are joined to reach this value and the rest of pixels form the next slice. This introduces a visualization delay in real-time applications, but it is always less than 32 pixels (the time delay depends on the device and the sweep speed). Time between slices encoding is always lower than 0.3 seconds for the ultrasound region. As each slice is coded separately, if an error occurs in the transmission it only affects to the quality of one slice. That is an important advantage compared with the MPEG-4 codecs that spread the error along the image and through several frames.

- **Image**. The codecs for the images are the same as slice codecs for the sweep video, 2-D SPIHT for image without color and 2-D CSPIHT for color image.

- **ECG Signal**. In [125] a real-time automatic ECG coding method that guarantees signal interpretation quality was designed. This coding method is based on 1-D SPIHT. It is the proposed solution in case of compress the ECG as signal.

- **Audio**. The sound has to be encoded with an audio codec for real time, such as Opus [126]. Opus is a highly versatile audio codec. Opus can handle a wide range of audio applications, but it is particularly suitable for interactive real-time applications over the Internet due to its low delay (22.5 ms by default). It can scale from low bit-rate narrowband speech to very high quality. Opus supports constant and variable bitrate encoding from 6 kbit/s to 510 kbit/s, frame sizes from 2.5 ms to 60 ms, and certain sampling rates from 8 kHz (with 4 kHz bandwidth) to 48 kHz (with 20 kHz bandwidth).

- **Text**. The text has to be encoded with a text codec, such as the ASCII 8-bit Unicode Transformation Format (UTF-8) [127]. The text is lossless coded.

In order to guarantee clinical quality of the encoded regions, a clinical evaluation of the ROI regions is necessary.

3.6 Results and Discussion for Compression for Real-Time Transmission

Previous results of clinical quality are not available for the proposed compression approach for real-time transmission of echocardiograms. Thus, a clinical evaluation following the methodology proposed in Section 3.2 was carried out for the compression method for real-time echocardiogram

Table 3.8: *Codec for every data type for compression for real-time transmission.*

Data Type	Codec	
	Black&White	Color
Video	3-D SPIHT [62]	
Sweep video	2-D SPIHT [41]	2-D CSPIHT [69]
Image	2-D SPIHT [41]	2-D CSPIHT [69]
ECG Signal	ECG SPIHT [125]	-
Audio	Opus [126]	-
Text	UTF-8 [127]	-

transmission and for different transmission rates. A precise evaluation of the real degradation in the compressed echocardiogram and a recommendation for the echocardiogram compression are provided. The recommended rates are compared with the transmission rates used in previous systems.

3.6.1 Evaluation Setup for Compression Recommendations

The database described in 3.1.2.1 has been used to carry out the evaluation described in the previous Section 3.2. Since the ultrasound region is the region that contains the most relevant clinical information, the evaluation has been carried out only for this region. The resolution of the ultrasound region for the different devices were: 668x496 for the Sonosite, 644x488 for the Envisor, and 634x462 for the IE33. Since not all the session revealed clinical information and much of the time was taken up with the process of measuring cardiac features (removed from the original video to avoid providing hints for later evaluations), the sessions were edited so as to remove these parts. The resultant duration of each video mode was approximately 1 min in the B and color Doppler modes and 30 s in the M and pulsed/continuous Doppler modes. Each session contained at least four videos of each mode and up to ten videos, having a total of 61 videos for the B mode, 63 for the color Doppler mode, 37 for the M mode and 39 for continuous/pulsed Doppler mode. The number of videos per device and mode are shown in Table 3.4. Audio was excluded from this study.

Three cardiologists experts in echocardiography interpretation, participated in the clinical evaluation. In order to evaluate the intra-observer variability, three sessions were evaluated twice for both tests and by the three cardiologists. The repeated sessions corresponded to different devices and rates. For the semi-blind tests, seven transmission rates were evaluated: 100, 150, 200, 250, 300, 400 and 500 kbps for the B and color Doppler modes, and 10, 15, 20, 25, 30, 40 and 50 kbps for the M and pulsed/continuous Doppler modes. These rates were chosen because their PSNR are considered adequate from the point of view of suitable clinical quality. In order to simplify the evaluation process, each pair of transmission rates and sessions was evaluated by two cardiologists. Thus, each cardiologist evaluated six sessions for each rate instead of nine. The total number of semi-blind tests carried out by each cardiologist was 180 (seven rates and six sessions, plus the three repeated sessions and 4 modes per session making a total of 180 tests). In each visualization the original and the compressed videos were visualized at the same time. Each cardiologist saw only two or three sessions (the four modes) per day so that this first evaluation lasted for approximately one month. Regarding the blind test, the three cardiologists evaluated two rates and the nine original sessions. Hence, the total number of blind tests carried out by each cardiologist was 120 (two rates plus the original session and nine sessions plus the three repeated sessions and 4 modes for session, making a total of 120 tests). Each cardiologist assessed a maximum of two sessions per day and six per week so that the evaluation took approximately two months.

The evaluation in [56] took two months too, but only four transmission rates were evaluated versus seven in the proposed evaluation. Furthermore, the second test, which is very burdensome, is only carried out for two transmission rates in the proposed evaluation versus four in [56].

3.6.2 Results for Compression for Real-Time Transmission

The results for the semi-blind test are shown in Tables 3.9-3.12, one table for each mode (B, color Doppler, M, and pulsed/continuous Doppler, respectively). The CDI_{SB} values (mean $\pm$ standard deviation of scores obtained from two cardiologists) for the nine patients and the three devices at different transmission rates are listed. The CDI_{SB} rated as inadequate are shown in dark gray, the CDI_{SB} with unacceptable quality but for which the same diagnosis is possible in light gray, and the CDI_{SB} with acceptable quality and which the same diagnosis is possible in white. As it has been described in the evaluation methodology, Section 3.2, the two rates selected are between the two latter ranges (light gray and white). Therefore, the highest selected transmission rate is the lowest rate with all the CDI_{SB} in white, and the lowest selected transmission rate is the rate immediately inferior. The two transmission rates selected for each mode appear in Tables 3.13. There is a clear effect of the transmission rate on the CDI_{SB} values for all the modes. The higher the transmission rate, the lower the CDI. Analysis of variance (ANOVA) confirmed that the CDI_{SB} values obtained for the different rates show significant differences ($p<0.05$) for all the modes.

Table 3.9: *Semi-blind test: CDI values for the B mode.*

Patient	Transmission rate (Kbps)						
	100	**150**	**200**	**250**	**300**	**400**	**500**
1	0.85± 0.07	0.25± 0.07	0.20± 0.00	0.20± 0.00	0.20± 0.00	0.20± 0.00	0.05± 0.07
2	0.90± 0.00	0.80± 0.00	0.25± 0.07	0.15± 0.07	0.15± 0.07	0.15± 0.07	0.10± 0.00
3	0.80± 0.14	0.75± 0.07	0.20± 0.00	0.20± 0.00	0.20± 0.00	0.15± 0.07	0.10± 0.00
4	0.25± 0.07	0.20± 0.14	0.20± 0.14	0.15± 0.07	0.10± 0.00	0.15± 0.07	0.05± 0.07
5	0.35± 0.07	0.20± 0.00	0.15± 0.07	0.10± 0.00	0.15± 0.07	0.15± 0.07	0.05± 0.07
6	0.85± 0.07	0.25± 0.07	0.15± 0.07	0.20± 0.00	0.15± 0.07	0.05± 0.07	0.05± 0.07
7	0.25± 0.07	0.15± 0.07	0.15± 0.07	0.15± 0.07	0.10± 0.00	0.15± 0.07	0.05± 0.07
8	0.30± 0.00	0.25± 0.07	0.15± 0.07	0.15± 0.07	0.20± 0.00	0.15± 0.07	0.05± 0.07
9	0.25± 0.07	0.20± 0.00	0.10± 0.00	0.15± 0.07	0.15± 0.07	0.05± 0.07	0.05± 0.07

Table 3.10: *Semi-blind test: CDI values for the color Doppler mode.*

Patient	**Transmission rate (Kbps)**						
	100	**150**	**200**	**250**	**300**	**400**	**500**
1	0.25± 0.07	0.20± 0.14	0.15± 0.07	0.15± 0.07	0.10± 0.00	0.15± 0.07	0.05± 0.07
2	0.25± 0.07	0.15± 0.07	0.15± 0.07	0.10± 0.00	0.15± 0.07	0.15± 0.07	0.05± 0.07
3	0.20± 0.00	0.20± 0.07	0.10± 0.00	0.15± 0.07	0.15± 0.07	0.10± 0.00	0.10± 0.00
4	0.25± 0.07	0.25± 0.07	0.20± 0.00	0.15± 0.07	0.10± 0.00	0.15± 0.07	0.05± 0.07
5	0.20± 0.00	0.20± 0.00	0.15± 0.07	0.10± 0.00	0.15± 0.07	0.10± 0.14	0.05± 0.07
6	0.75± 0.07	0.25± 0.07	0.05± 0.07	0.10± 0.00	0.05± 0.07	0.00± 0.00	0.00± 0.00
7	0.20± 0.00	0.15± 0.07	0.15± 0.07	0.15± 0.07	0.05± 0.07	0.10± 0.00	0.05± 0.07
8	0.25± 0.07	0.25± 0.07	0.15± 0.07	0.15± 0.07	0.15± 0.07	0.10± 0.14	0.00± 0.00
9	0.25± 0.07	0.25± 0.07	0.15± 0.07	0.15± 0.07	0.15± 0.07	0.05± 0.07	0.10± 0.00

Table 3.11: *Semi-blind test: CDI values for the M mode.*

Patient	Transmission rate (Kbps)						
	10	**15**	**20**	**25**	**30**	**40**	**50**
1	0.90± 0.00	0.80± 0.00	0.35± 0.07	0.25± 0.07	0.25± 0.07	0.20± 0.00	0.05± 0.07
2	0.90± 0.00	0.85± 0.07	0.25± 0.07	0.25± 0.07	0.25± 0.07	0.20± 0.00	0.10± 0.00
3	0.85± 0.07	0.30± 0.00	0.25± 0.07	0.25± 0.07	0.20± 0.00	0.20± 0.00	0.05± 0.07
4	0.30± 0.00	0.25± 0.07	0.20± 0.14	0.20± 0.14	0.00± 0.00	0.10± 0.14	0.05± 0.07
5	0.80± 0.00	0.20± 0.00	0.15± 0.07	0.10± 0.00	0.10± 0.14	0.10± 0.14	0.00± 0.00
6	0.80± 0.00	0.25± 0.07	0.25± 0.07	0.25± 0.07	0.10± 0.14	0.10± 0.14	0.05± 0.07
7	0.25± 0.07	0.25± 0.07	0.25± 0.07	0.15± 0.07	0.15± 0.07	0.15± 0.07	0.00± 0.00
8	0.85± 0.07	0.35± 0.07	0.10± 0.14	0.20± 0.14	0.20± 0.00	0.15± 0.07	0.00± 0.00
9	0.20± 0.00	0.20± 0.00	0.15± 0.07	0.15± 0.07	0.15± 0.07	0.10± 0.00	0.00± 0.00

Table 3.12: *Semi-blind test: CDI values for the pulsed/continuous Doppler mode.*

Patient	**Transmission rate (Kbps)**						
	10	**15**	**20**	**25**	**30**	**40**	**50**
1	0.85± 0.07	0.55± 0.35	0.30± 0.00	0.25± 0.07	0.25± 0.07	0.20± 0.00	0.00± 0.00
2	0.35± 0.07	0.30± 0.00	0.25± 0.07	0.20± 0.00	0.20± 0.14	0.10± 0.00	0.00± 0.00
3	0.80± 0.00	0.30± 0.00	0.25± 0.07	0.25± 0.00	0.20± 0.14	0.10± 0.00	0.00± 0.00
4	0.30± 0.00	0.25± 0.07	0.20± 0.14	0.10± 0.00	0.10± 0.14	0.10± 0.14	0.00± 0.00
5	0.30± 0.00	0.30± 0.00	0.10± 0.00	0.15± 0.07	0.05± 0.07	0.10± 0.14	0.00± 0.00
6	0.85± 0.07	0.30± 0.00	0.25± 0.07	0.20± 0.00	0.15± 0.07	0.05± 0.07	0.05± 0.07
7	0.80± 0.00	0.25± 0.07	0.20± 0.00	0.15± 0.07	0.15± 0.07	0.10± 0.14	0.00± 0.00
8	0.30± 0.00	0.30± 0.00	0.20± 0.14	0.15± 0.07	0.15± 0.07	0.10± 0.14	0.05± 0.07
9	0.85± 0.07	0.25± 0.07	0.25± 0.07	0.15± 0.07	0.15± 0.07	0.05± 0.07	0.00± 0.00

Table 3.13: *Selected transmission rates to be evaluated in the blind test.*

B	D	M	DP
200 kbps	150 kbps	30 kbps	30 kbps
250 kbps	200 kbps	40 kbps	40 kbps

The results for the blind test are shown in Table 3.14. The transmission rates shown in this table are those selected from the semi-blind test (Table 3.13). The first row shows the operation mode and the second the target transmission rate. The other cells contain the CDI values (mean $\pm$ standard deviation of scores obtained from the three cardiologists) for the nine patients and the three devices at different transmission rates. The CDI values higher than the maximum acceptable (0.25) are colored in gray.

In order to test the intra-observer variability each cardiologist evaluated three sessions twice for both tests and the four modes. ANOVA study was performed using Matlab to see whether there were significant differences between the repeated measures. There are four different statistics to calculate the significance value. They are Pillai's trace, Hotelling-Lawley's trace, Wilk's lambda and Roy's largest root. For the semi-blind test two studies were performed. One for the D values and other for the C values. The study was divided because it is very important that the D does not change since it is indicating whether the diagnostic is valid or not. However it is not so important a little variation in the subjective measurement of the similarity. For the blind test the reasoning is the same. We separated the Q quality scores of the I interpretations. For both tests, the D and I values were the same for the repeated tests and the three cardiologists, having a significant value of 1 in all the cases and statistics. For the C and Q values the significant values were in all the cases higher than 0.8, thus the null hypothesis was tested, there are not significant differences between the repeated measurements.

Table 3.14: *Blind test: CDI values for the four modes.*

Patient	Bit rate (Kbps) per mode							
	B		Doppler		M		P/C Doppler	
	200	250	150	200	30	40	30	40
1	0.15± 0.03	0.15± 0.03	0.16± 0.12	0.13± 0.04	0.15± 0.06	0.15± 0.06	0.19± 0.10	0.16± 0.02
2	0.15± 0.06	0.08± 0.05	0.14± 0.06	0.11± 0.06	0.31± 0.20	0.16± 0.02	0.25± 0.18	0.13± 0.04
3	0.13± 0.05	0.13± 0.03	0.21± 0.09	0.12± 0.03	0.14± 0.07	0.11± 0.02	0.14± 0.07	0.08± 0.02
4	0.16± 0.14	0.11± 0.08	0.20± 0.12	0.14± 0.04	0.09± 0.14	0.02± 0.04	0.10± 0.11	0.10± 0.10
5	0.18± 0.11	0.13± 0.05	0.19± 0.05	0.16± 0.07	0.13± 0.10	0.13± 0.10	0.10± 0.11	0.07± 0.04
6	0.15± 0.13	0.14± 0.04	0.20± 0.12	0.08± 0.09	0.13± 0.14	0.08± 0.12	0.14± 0.11	0.05± 0.08
7	0.14± 0.14	0.12± 0.07	0.19± 0.13	0.15± 0.07	0.12± 0.07	0.10± 0.07	0.13± 0.05	0.12± 0.10
8	0.19± 0.09	0.11± 0.10	0.25± 0.06	0.15± 0.10	0.16± 0.10	0.13± 0.08	0.16± 0.07	0.11± 0.13
9	0.15± 0.02	0.11± 0.08	0.20± 0.06	0.11± 0.07	0.14± 0.04	0.07± 0.04	0.14± 0.07	0.03± 0.04

3.6.3 Discussion for Compression for Real-Time Transmission

3.6.3.1 Compression Recommendations

As explained in Section 3.2, the CDI_{SB} values were considered to obtain the two transmission rates that were used in the blind test. The standard deviation of the CDI_{SB} values is very low for all the modes, indicating low inter-observer variability. This is because the cardiologists have the same opinion about the cases in which both videos have the same diagnosis and similar opinions in all the cases about the similarity of the videos, but it is not always the same. The higher the transmission rate, the higher the visible quality. This is consistent with the clinical results shown in Tables 3.9-3.12.

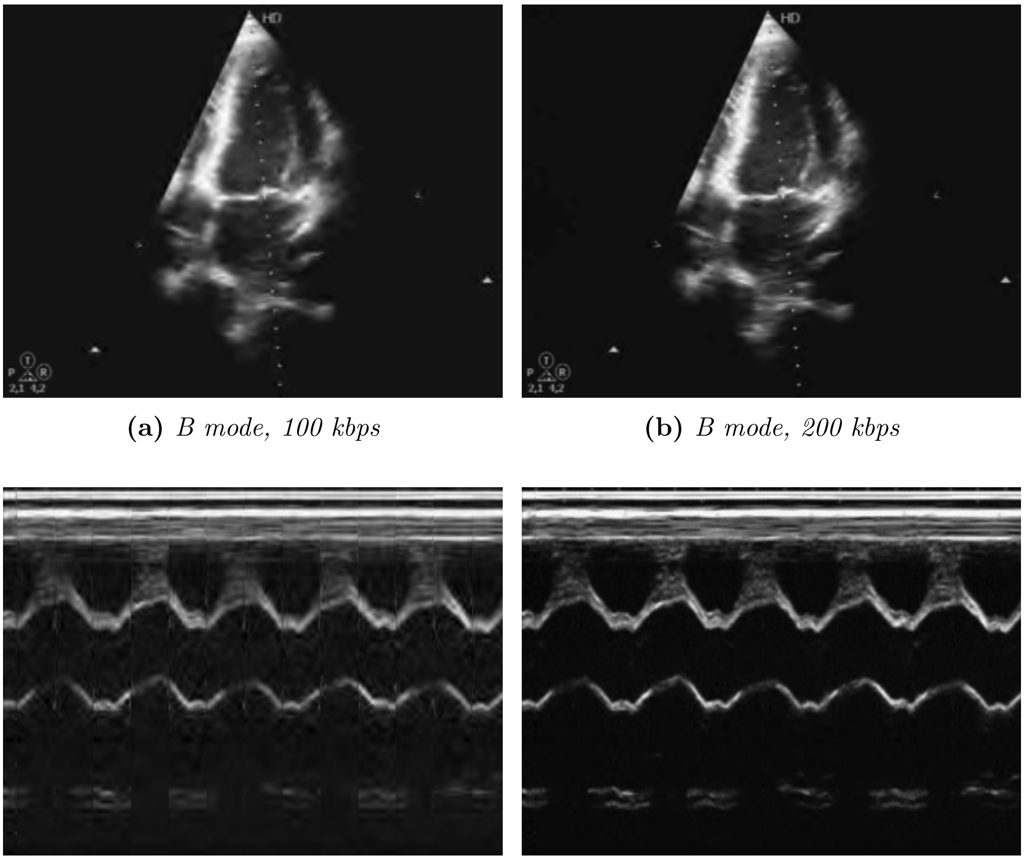
Now the recommended transmission rates for the proposed compression method are provided, as explained in Section 3.2. Table 3.14 shows the CDI values for the two selected transmission rates for each mode. It can be seen that the CDI values for both transmission rates are very similar. This is because very good results were obtained for both rates and the diagnosis was practically identical to the original in most cases. Thus the CDI values are more affected by the quality score. In general, the standard deviation of the CDI values is lower for the highest transmission rates. The reason is that for the highest rates the image quality is clearly good whereas for the lowest rates it is a little less clear and the cardiologists have slightly different opinions. In Table 3.14, we can see in gray the CDI values which are higher than the maximum acceptable value (0.25). The selected transmission rates are the lowest rates with all the CDI values below 0.25. This constraint leads to the recommendations for transmission rates shown in Table 3.15. If we look at the CDI_{SB} values of the semi-blind test (Tables 3.9-3.12) for the recommended transmission rates, we can see that all the values are equal or lower than 0.2 except for one mode. Hence for all the modes except for the B mode a recommended transmission rate with only the semi-blind test can be given, being the recommended transmission rate that with all the CDI_{SB} values equal or lower than 0.2. For the B mode the blind test is necessary, because this mode contains a lot of information and it is more difficult to evaluate.

In Figure 3.15 two different images compressed at different transmission rates are shown. Figures 3.15b and 3.15d show images compressed at the recommend rate for two modes. The images with lower transmission rates than the recommended can present some artifacts, such as a loss of clarity and sharpness in the edges and structures that may cause clinicians to provide an erroneous diagnostic (see Figures 3.15a and 3.15c). For the M mode and the lowest transmission rate, we can see a loss of clarity in the slice edges. This is due to the compression method for the sweep modes, that compresses each slice separately. However, this artifact is no visible for the recommended rate.

In order to test the inter-observer variability, a study similar to the intra-observer variability has been performed. The variance study has been performed for the selected rate and every mode proving the results that there are not significant differences between the cardiologist measurements.

It is very important to comment that the transmission recommendations are only applicable to the proposed compression method and a general echocardiogram examination. The results obtained are independent of the devices and diagnosis and therefore the recommendations may be generalized to any echocardiogram device and any diagnosis.

Tabla 3.16 shows the codecs and compression recommendations for each region depending on the data type. The ultrasound regions are compressed with the recommended rates previously



(a) *B mode, 100 kbps*

(b) *B mode, 200 kbps*

(c) *M mode, 10 kbps*

(d) *M mode, 40 kbps*

Figure 3.15: *Images of the B and M modes compressed at different rates with the proposed method.*

assessed. The rest of images are compressed with 0.5 bpp, which presents a good quality in medical image compression [37]. The ECG signal recommendation is in [125]. The auxiliary video, which has a small size, and sound are compressed with typical compression rates.

Table 3.15: *Recommended transmission rates per mode to obtain good clinical quality.*

B	**D**	**M**	**DP**
200 kbps	200 kbps	40 kbps	40 kbps

3.6.3.2 Codecs Comparison

Section 2.7 lists the resolution, codecs and transmission rates for the most relevant ultrasound video systems. It can be seen that the recommended transmission rates using the compression proposed in this Thesis are lower than those used in the other systems, even for the transmission rates with low resolution video. The best compression result for high video resolution was presented in [56]. The recommended transmission rates were 768 Kbps for the B and M mode, and 256 Kbps for the color Doppler and pulsed/continuous Doppler. Hence, the proposed method requires less than 26% of the amount of data for the B mode, 78% for the Color Doppler mode, 5.2% for the M mode, and 16% for pulsed/continuous Doppler.

We can conclude that the results obtained with the proposed technique clearly show an improved performance in terms of the transmission rate as compared with the other codecs presented in the literature for ultrasound videos (Xvid, H.264/AVC, Windows Media and HEVC). The low transmission rates are obtained mainly because the echocardiogram characteristics have been taken into account in the compression design and a clinical evaluation has been carried out in order to obtain the minimal recommended rates. This saving in the transmission rate will lead to better transmission performance.

3.7 Conclusions

The main objective of this Chapter is to achieve the efficient compression of echocardiograms. This overall objective has been divided into three: to develop a clinical evaluation methodology for transmission rate recommendations, to design a compression method to store echocardiograms and to design a compression method to transmit echocardiograms in real-time. The main conclusions relating to these three specific objectives are listed below:

- An evaluation methodology which is accurate but not very time consuming has been designed for compressed echocardiogram. The proposed method can be used to assess any video codec and to recommend a minimal transmission rate which guarantees clinical quality. For all the modes except the B mode it is only necessary to perform the semi-blind test because it gives the same result as the blind test. This evaluation methodology may be adapted to other ultrasound techniques or modes, or even to other medical image modalities.

Table 3.16: *Codecs and recommended transmission rates and bits per pixel for each type of region.*

Regions	Data type	Codec	Compression
Ultrasound	Video	3-D SPIHT [62]	200 Kbps
	Sweep video	2-D SPIHT [41], 2-D CSPIHT [69]	40 Kbps
ECG	Signal	ECG SPIHT [125]	500 bps
	Sweep video	2-D SPIHT [41], 2-D CSPIHT [69]	0.5 bpp
Auxiliary video	Video	3-D SPIHT [62]	100 Kbps
Auxiliary image	Image	2-D SPIHT [41], 2-D CSPIHT [69]	0.5 bpp
Sound	Audio	Opus [126]	10 Kbps
Text	Text	UTF-8 [127]	1-4 bytes per character

- An image compression format for storage purposes has been proposed that takes advantage of the segmentation facilities of the device to enhance the compression performance. This compression format is easily integrable in the acquisition device without adding complexity to devices that already incorporate the DICOM standard. The proposed method shows better results than compression without using regions for all the available dataset, having a compression gain ranging from 14 % up to 75 % for a typical bit rate. Although the compression results depend on the acquisition devices, how the image is displayed, and the compression quality, the compression ratios obtained for the DICOM file ranged from 19 to 41 without losing diagnostic information. Furthermore, a tool that makes conversions between different image formats has been developed. This tool allows interoperability between medical centers and devices, and also enables echocardiogramsto be stored with the proposed format, thus saving storage space. Options to edit the image (edit/add/remove regions) taking advantage of the proposed format based on regions have been also included.

- An echocardiogram compression method for real-time transmission based on regions and visualization modes has been designed. Codecs and transmission rate recommendations have been given for each region. Since the ultrasound region contains the most important information from a medical point of view, a comprehensive evaluation has been preformed for this region. Minimum transmission rates have been recommended for each mode in order to guarantee suitable clinical quality for the transmission and storage of echocardiogram videos using the proposed technique. The recommended transmission rates for the ultrasound regions are the following: 200 Kbps for the 2-D and the color Doppler modes, and 40 Kbps for the M and the pulsed/continuous Doppler modes. These are very good results in terms of bandwidth use, especially for the M and pulsed/continuous Doppler modes that have been obtained thanks to the fact that the compression method takes into account the stationary characteristics of the sweep modes and only a thin slice is compressed for each frame. The compression recommendation for the other regions has also been provided. Moreover, these results make possible the transmission of echocardiogram videos over 3G wireless networks and beyond. The recommended transmission rates for previous ultrasound transmission systems are considerably higher than those given for the proposed method which allows better transmission results.

Chapter 4

Echocardiogram Transmission in Real-time

This Chapter deals with the second part of the tele-echocardiography systems (see Figure 4.1): transmission and display. Once the compression recommendations have been established, it is necessary to guarantee that the echocardiogram is received without diagnostic information being lost. If the transmission is completed without errors, and consequently the whole echocardiogram is visualized with the recommended transmission rate, the same diagnosis as that of the original echocardiogram is guaranteed. However, if errors occur in the transmission process, some echocardiogram parts will be visualized with a transmission rate lower than the recommended rate. In that case the diagnosis may be possible or not. For this reason, an evaluation for the echocardiogram display recommendations has been designed and carried out for the echocardiograms compressed with the method proposed in this Thesis for real-time transmission. The display recommendation allows us to know if the echocardiogram is visualized without losing diagnostic information for any transmission conditions instead of having to carry out different assessments for each channel condition. In addition, a protocol is required for the real-time end-to-end transmission of echocardiograms over IP that defines how to transmit each encoded region and the synchronism between them. In the case of WiMAX channels, which introduce packet losses, an error control method is also required for the regions with relevant clinical information. The echocardiogram database and characteristics for real-time transmission purposes used in this Chapter were described in the previous Chapter, Section 3.1. This Chapter is organized as follows. Section 4.1 describes the evaluation methodology for display recommendations of medical images after transmission, and specifically for echocardiograms that have been encoded with the method proposed in this Thesis for real-time transmission. Section 4.2 gives the display recommendations for the echocardiogram after compression with the proposed method and recommended transmission rates for real-time transmission as listed in Chapter 3. Section 4.3 describes the echocardiogram transmission protocol. Section 4.4 discusses the error control method and the configuration for each type of region. Section 4.5 provides the results and discussion relating to simulations of real-time echocardiogram transmission over WiMAX channels carried out both with the compression and transmission methods proposed in this Thesis and without them in order to evaluate the improvements of each proposed method. Finally, the conclusions of this Chapter are given in Section 4.6.

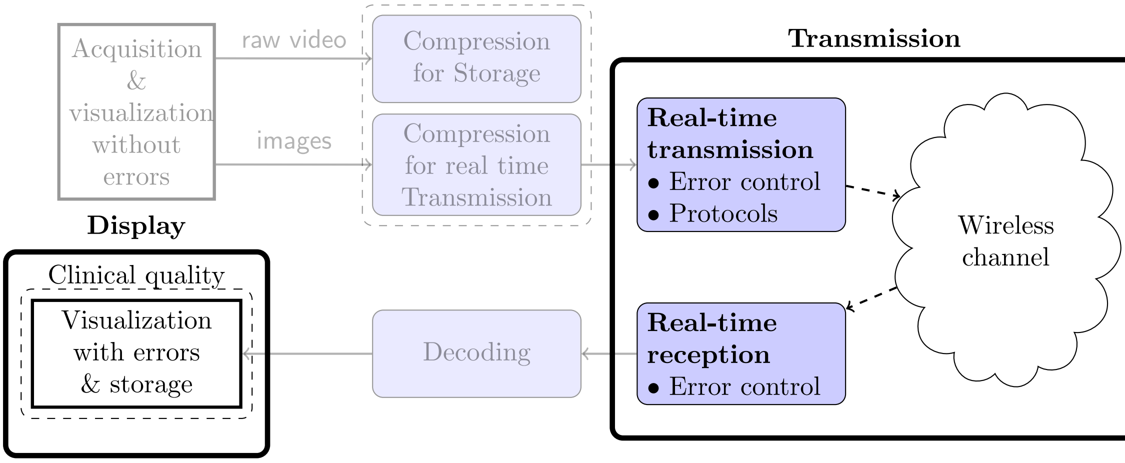


Figure 4.1: *Tele-echocardiography system structure followed in this Thesis: transmission and visualization parts.*

4.1 Clinical Evaluation Methodology for Display Recommendations

This section presents an evaluation methodology to give recommendations for the display of echocardiograms after transmission. The starting point of the evaluation is the recommendations already given for the compression. If errors occur in the transmission, some parts of the echocardiogram are visualized with a transmission rate lower than the recommended rate, but this does not mean that the diagnosis is not possible. The maximum acceptable time during which the echocardiogram is visualized with a transmission rate lower than the recommended rate has to be determined. In order to give recommendations that are independent of the transmission channel, the worst case scenario has to be evaluated. The evaluation methodology has been designed taking into account the compression approaches followed in this Thesis for real-time transmission of the ultrasound regions. Different evaluations are proposed for each visualization mode:

- **2-D modes**. The evaluation establishes the maximum percentage of time that the echocardiogram can be visualized with a transmission rate lower than the recommended rate. Different percentages of time and transmission rates have to be assessed for a 2-D video.
- **Sweep modes**. For the sweep modes, the echocardiogram is compressed by slices. The important part for the diagnosis is when the sweep is stopped and the screen is filled with slices. For this reason, instead of evaluating the percentage of time with a lower than recommended transmission rate, the number of slices that can be visualized with a lower transmission rate is evaluated. Different numbers of slices and transmission rates have to be assessed for an image. This image is that used by the cardiologist to perform the diagnosis.

The evaluation for the two types of visualization modes consists of a semi-blind test. In this case a blind test is not necessary because there is already adequate clinical quality resulting from the previously recommended transmission rates. The objective of this test is to determine whether or not the cardiologist would be able to make the same diagnosis with the video or images compressed at the transmission rate recommended in Chapter 3 and with the video or image with some parts compressed with a lower transmission rate due to transmission errors. This test consists of giving an opinion about the similarity between the two compared videos or images (see Figure 4.2). It is also established whether the diagnosis would be the same with both videos and images. The same diagnosis is not possible if the mark is lower than 3; otherwise, the same diagnosis is possible with either videos or images.

1. Measure of similarity between the two videos or images.

1: very different	2: different	3: acceptable	4: similar	5: identical

Figure 4.2: *Semi-blind test: comparison of compressed echocardiogram with the recommended transmission rates.*

A clinical distortion index (CDI) is calculated in order to have an estimation that directly reflects if the evaluated video or image having some parts with a lower than recommended transmission rate is of sufficient clinical quality for an adequate diagnosis. A CDI is calculated for each echocardigram operation mode, video or image, with different percentages of time of the total time or numbers of slices with lower than recommended transmission rates. This is defined as

$$CDI \;=\; min\{D_i\} \tag{4.1}$$

where D_i is the measurement of the test in Figure 4.2 by different cardiologists and images or videos compressed with the conditions to be evaluated. The minimum D_i value is selected, because in the event that one of the evaluated visualizations is not correct, the conditions are not suitable for a valid diagnosis.

The CDI value can be divided into two quality ranges:

- $CDI < 3$: the same diagnosis is not possible and the quality is not sufficient for the evaluated videos or images.
- $CDI \geq 3$: the same diagnosis is possible and the quality is sufficient for all the evaluated videos or images.

With the calculated CDI value, the maximum percentages of time or maximum number of slices with the different transmission rates lower than the recommended rate will be chosen for each mode. These values are the highest with CDI values higher or equal to 3. This is because the cardiologists have decided that the minimum acceptable value is 3.

This evaluation methodology is valid for compression based on visualization modes. However, if no visualization modes are distinguished in the compression process, the methodology for 2-D modes can be applied for all other modes.

These recommendations determine whether each fragment of the echocardiogram that corresponds with a change of operation mode has sufficient clinical quality. Each fragment has to be checked separately. If all the fragments are valid, the whole echocardiogram is valid. In the event that a mode is not suitably visualized, the whole mode has to be transmitted again, delaying the diagnosis process.

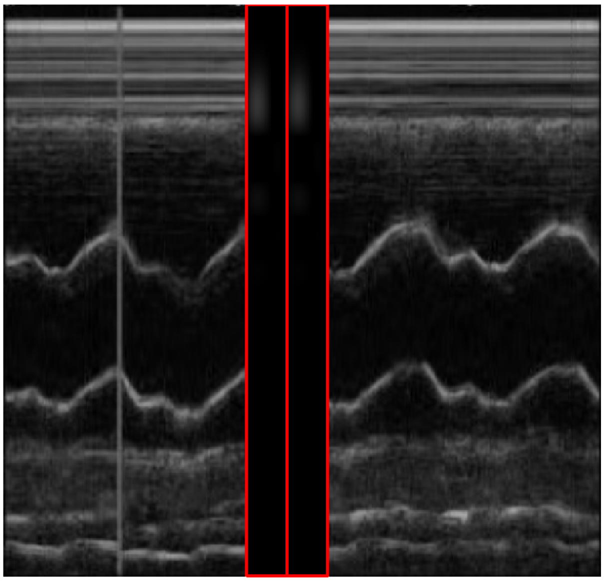
4.2 Results and Discussion for Display Recommendations

The clinical evaluation proposed in Section 4.1 was carried out for the echocardiograms that are compressed with the method proposed in this Thesis and the recommended transmission rates listed in Chapter 3. Display recommendations are provided for each echocardiogram operation mode independently of the transmission channel. Thus, the display recommendations allow us to determine whether the echocardiogram fragments that correspond to a mode have adequate clinical quality without needing to perform an evaluation for each transmitted echocardiogram.

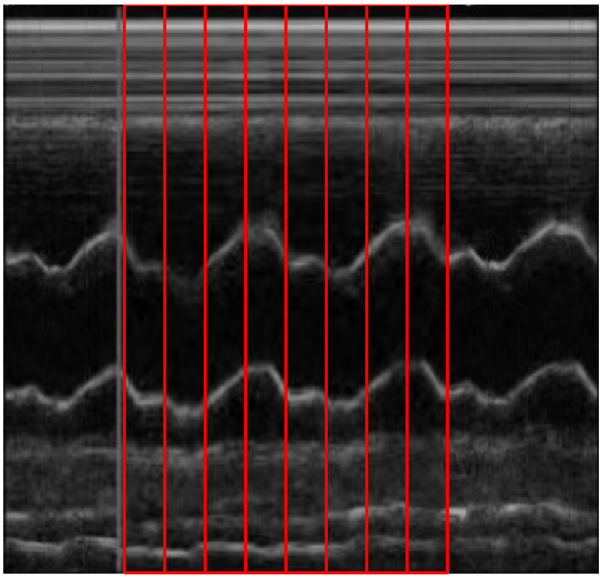
4.2.1 Evaluation Setup for Display Recommendations

The database described in Section 3.1.2.1 has been used to carry out the evaluation. Since the ultrasound region is the region that contains the most relevant clinical information, the evaluation has been carried out for this region only. The echocardiograms were edited so as to remove the part without clinical information. The number of videos per device and mode are shown in Table 3.4. Each session contains videos of about 1 minute of duration for the 2-D modes and one image per video for the sweep modes.

Two cardiologists expert in echocardiography interpretation participated in the clinical evaluation. In order to evaluate the intra-observer variability, three sessions were evaluated twice for both tests and by the two cardiologists. The repeated sessions corresponded to different devices, percentages of time and numbers of slices with transmission rates lower than recommended, and transmission rates lower than recommended. The echocardiograms were compressed with the proposed techniques and recommended transmission rates for each mode for the ultrasound region (200 kbps for the 2-D modes and 40 kbps for the sweep modes). We have to take into account that an error affects to all the bits encoded together, 16 frames for 2-D modes and 1 slice for the sweep modes. As the compression algorithms are embeddedness, for the bits that are encoded together, only the first error affects. When an error occurs the rest of bits are ignored and the region is decompressed with the received previous bits. For the sweep modes, the numbers of slices evaluated with transmission rates inferior to the recommend rate were 2, 4, 6, 7, 8, 9 and 10. The transmission rates were 20 kbps (50 % of the recommended transmission rate) and 0 kbps. The 0 kbps transmission rate simulates the case in which no packets are received. The slices with the lowest transmission rate were located in the middle of the screen to simulate the worst case, with the worst clinical quality in order to give recommendations that do not depend on the transmission channel. In Figure 4.3 two samples of images that have been evaluated for a M mode are illustrated. Figure 4.3a shows an example of an image with two slices with a transmission rate of 0 kbps and the rest of the slices with 40 kbps (recommended transmission rate). Figure 4.3b shows an example of an image with eight slices with a transmission rate of 20 kbps (50 % of the recommended rate) and the rest of the slices with a transmission rate of 40 kbps. The total number of slices for the whole screen depends on the device but is about 20 for the available database. For the 2-D modes, the evaluated percentages of time with transmission rates lower than the recommended rate were 5%, 20% and 50%. The transmission rates were 160 kbps (80 % of the recommended transmission rate), 100 kbps (50 % of the recommended transmission rate) and 0 kbps. The bandwidth with inferior quality is continuous and is located in the middle of each evaluated video to simulate the worst possible case. This is so that the recommendations do not depend on the transmission channel. A continuous distribution of the bandwidth with a lower transmission rate than the recommended rate is worse than a burst distribution. Figure 4.4 shows the bandwidth distributions over time for a video with 20 % of the time with 160 kbps (drawn in dark blue) and for another video with 5 % of the time with 0 kbps (drawn in sky blue). Each cardiologist saw only one visualization quality of the same mode per day so that the evaluation took approximately 17 days.



(a) *Image with two slices (in red) with transmission rate of 0 kbps.*



(b) *Image with eight slices (in red) with transmission rate of 20 kbps.*

Figure 4.3: *Samples of M mode images for the evaluation of the display recommendations.*

4.2.2 Results for Display Recommendations

The results of the test are shown in Tables 4.1 and 4.2, one table for each type of visualization mode (2-D and sweep modes, respectively). The CDI values rated as adequate are shown in dark gray and the CDI values with inadequate quality are in white. The maximum acceptable number of slices and percentages of time that the echocardiogram is visualized with a transmission rate lower than the recommended rate is the highest with a CDI value higher or equal to 3. There is a clear effect of the transmission rate, percentage of time and number of slices on the CDI values. The higher the transmission rate, the higher the CDI value. The higher the percentage of time and number of slices, the lower the CDI value.

As regards inter-observer variability, both cardiologists were of the same opinion as to whether or not the video or image had adequate clinical quality (CDI value higher or equal to 3 or lower than 3). However, they had slightly differing opinions about the quality (in the same CDI range, but different values). The intra-observed evaluations show that there are no diagnostic differences between the repeated measurements.

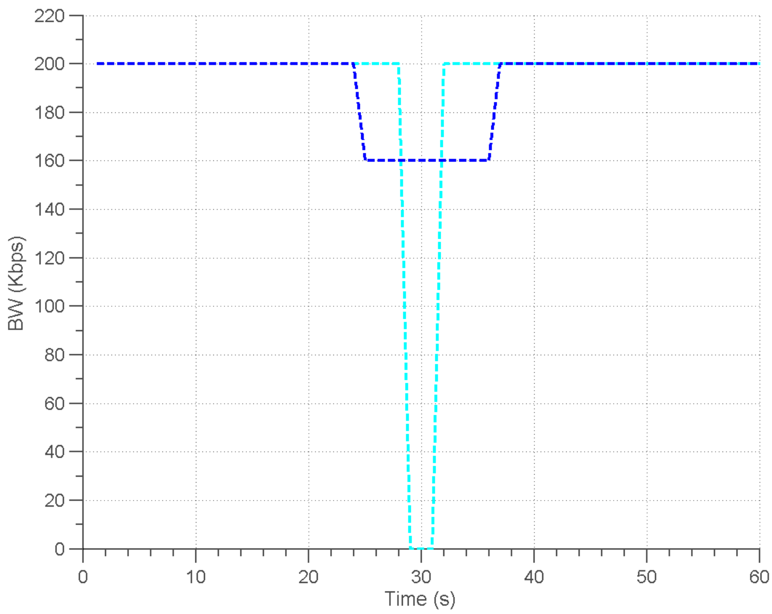


Figure 4.4: *Samples of visualized bandwidth distribution for the evaluation of the display recommendations of a 2-D mode video of 1 minute duration. The bandwidth with 20 % of the time with 160 kbps is shown in dark blue. The bandwidth with 5 % of the time with 0 kbps is shown in sky blue.*

Table 4.1: *CDI values for the 2-D modes.*

% time	**Transmission rate (kbps)**					
	B mode			**Doppler mode**		
	160	**100**	**0**	**160**	**100**	**0**
5	3	3	3	4	3	3
20	3	2	2	3	3	2
50	2	2	1	2	2	1

4.2.3 Discussion for Display Recommendations

Given the results in Tables 4.1 and 4.2 and the methodology in Section 4.1, the following display recommendations are given for each mode:

- B mode.
  - Up to 5% of the time the B mode can be visualized with any transmission rate.
  - Up to 20% of the time the B mode can be visualized with a transmission rate of 160 kbps or higher.
- Color Doppler mode.
  - Up to 5% of the time the color Doppler mode can be visualized with any transmission rate.

Table 4.2: *CDI values for the sweep modes.*

# slices	Transmission rate (kbps)			
	M mode		DP/DC mode	
	20	**0**	**20**	**0**
2	3	3	4	3
4	3	1	3	2
6	3	1	3	2
7	3	1	3	2
8	3	2	3	2
9	2	2	3	1
10	2	1	2	1

  - Up to 20% of the time the color Doppler mode can be visualized with a transmission rate of 100 kbps or higher.

- M mode.
  - Up to 2 slices can be visualized with any transmission rate (see Figure 4.3a).
  - Up to 8 slices can be visualized with a transmission rate of 20 kbps or higher (see Figure 4.3b).
- Pulsed/continuous Doppler mode.
  - Up to 2 slices can be visualized with any transmission rate.
  - Up to 9 slices can be visualized with a transmission rate of 20 kbps or higher.

There are two display recommendations for every mode. The display recommendations listed first, corresponding to any transmission rate, are valid for any video codec for the 2-D modes and for any image codec for the sweep modes if the recommended transmission rates for the compression have being previously defined. However, the second display recommendations are only valid for the proposed compression method for the ultrasound regions. In the event of using a compression without distinguishing modes and with 3-D SPIHT for all the modes, the visualization recommendations will be the same as the B mode for all the modes. It is important to emphasize that in order to determine whether the echocardiogram is visualized with adequate diagnostic information, each operation mode fragment has to be checked separately.

4.3 Protocol for Echocardiogram Transmission in Real-time

A first approach to the problem of echocardiogram transmission was described in [128]. An enhanced protocol for real-time transmission was proposed. Although good transmission results were obtained, the protocol did not make use of the visualization characteristics and segmentation of echocardiograms. All the visualization modes were encoded and transmitted in the same way, and as a result the maximum transmission rates of the modes were used. After studying new improvements, a similar protocol was proposed in [129] but this time encoding the echocardiogram by visualization modes. A lower transmission rate was used in the sweep modes and consequently better transmission performance was obtained with this technique. However, it did not incorporate region segmentation.

Finally, after taking into account all the research experience and know-how developed over the years together with previous versions of the protocol [128,129], we propose an Echocardiogram Transmission Protocol (ETP) application protocol for real-time and end-to-end transmission over IP of echocardiograms encoded by visualization modes and regions. This is described in the following sections.

4.3.1 ETP Overview

ETP takes advantage of the visualization characteristics and the facilities of segmentation that the acquisition devices already incorporate. Since each region is encoded separately, it is proposed to send each region separately to obtain better results in the transmission process. Furthermore, this way of encoding and sending each region separately allows to allocate different degrees of priority and to use either a TCP or a UDP transport layer protocol for each region. Each echocardiogram device has its own distribution of regions and even the distribution can change during transmission. Therefore, not only is it necessary to send the encoded regions, but it is also necessary to send control information such as the configuration of the regions or the synchronism between them. There are thus, two types of packets: control and data packets.

- The **control packets** contain the control information. This consists of configuration parameters such as the image size and region parameters such as the type of region, its size and its position. It also contains information to establish the data connections. The configuration is codified with an XML file. XML is proposed because of the advantages mentioned in the previous Chapter 3 for the compression for storage. The encoded text and auxiliary images are also included in these packets because they do not need to be synchronized with the rest of the regions (see Table 3.6).
- The **data packets** contain the encoded information of each region. There are different flows of packets for each region that need to be synchronized (see Table 3.6). Each region is encoded according to its data type (see Table 3.16), image, video, sweep video, sound or text. The specific codec is defined in the configuration for each region although the default recommended codecs and the used in this Thesis are shown in Table 3.16.

Figure 4.5 shows the protocol flow for a Sonosite echocardiogram, where for each mode only the ultrasound region has data connection. The control packets are sent at the beginning, before the data packets are sent, to set the initial configuration and to open the data connections. They are also occasionally sent when some text, auxiliary image or region change. When the data connections have been established, the data packets of the activated regions are sent. The ultrasound region indicates the mode that is activated in each moment, and consequently the activated regions that correspond with the activated mode. Furthermore, the ultrasound part can be stopped for measurements to be taken, during which time no data packets are sent. A control packet is also sent to finish the transmission, the connections are closed and the echocardiogram display ends.

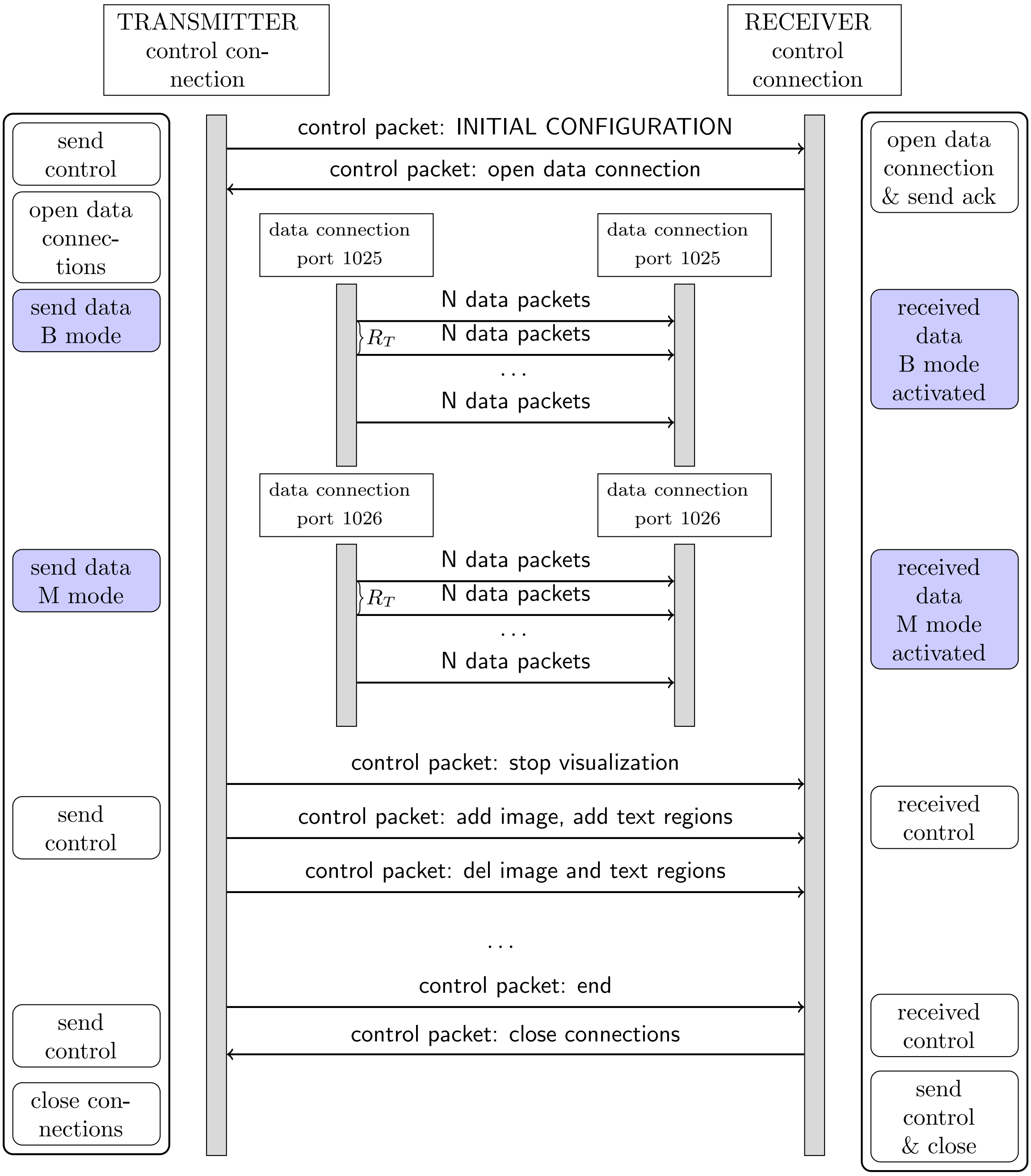


Figure 4.5: *Protocol flow diagram for the Sonosite device echocardiogram transmission. The control packets (configuration and confirmation) are in white and the data packets in blue. N is the number of packets sent each resolution time and R_T is the resolution time.*

TCP is proposed for sending the control packets because it is important to guarantee their reliability while their arrival time is less important. The text has to be protected so that it can only be accessed by authorized sanitary staff. An easy way to protect this information is by protecting all the configuration packets in which the information is contained. The TLS protocol has been proposed because it is the most common choice for secure communications in medical standards and it does not introduce overhead to the transmitted packets.

UDP is proposed for sending the data packets since the flow of packets is high and it is well known that UDP works well for real time transmission of medical video flows. Thus, for each region a UDP connection is open in the port indicated in the control packets. In order to support the UDP protocol, considering that UDP does not implement reliability, an error control method is essential in clinical videos in order not to lose diagnostic information. The appropriate reliability method depends on the channel used for the transmission and also on the region flow and clinical relevance. The used error control method will be indicated in the control packets for each region allowing the use of the ETP protocol for any channel. In Section 4.4 an error control method is proposed for the transmission of echocardiograms over wireless networks.

In order to decrease header overheads, reduce packet loss and increase security over noisy wireless links, ROHC can be used for both UDP [81] and TCP [82] transport layers. This standard compresses IP and UDP headers to just 3 bytes, and IP and TCP headers to just 10 bytes.

4.3.2 Control Packets Coding

There are two types of control packets: configuration and confirmation packets. The configuration packets, which the transmitter sends to the receiver, contain the global configuration (see Table 4.3), the configuration of each region that is common to all the regions (see Table 4.4), the particular configuration of each region and the coding of the regions that hardly changes (see Table 4.5), i.e. text and auxiliary images. The contents of the XML file are described in Tables 4.3-4.5. The syntax has to be defined, for example, with a DTD or an XML Schema file. The proposed DTD file is defined in Appendix A. An example of an XML file for the color Doppler and pulsed Doppler mode from Figure 4.7 and for regions with ID 1, 2, 3, 4, and 10 is shown in Table 4.6.

The XML file proposed for real-time transmission is not the same as that proposed for storage (see Section 3.3) because echocardiograms for storage and for real-time transmission purposes have different characteristics. However, as both echocardiogram types have the same distribution of regions, the XML files share some fields, for example the *region* and *tsize* fields. The echocardiogram can be stored at the same time that it is visualized by the cardiologist. The cardiologist chooses the relevant images for the diagnosis to be stored. The translation of the XML file to storage format is easy to perform since both XML files contain common information. Table 4.7 shows the XML files, real-time transmission configuration and storage format for the regions with ID 1 and 2 for the B and M modes of the Sonosite devices in Figure 3.3 and their equivalent. For the whole echocardiogram, it is necessary to add the other regions and the regions that correspond to the other modes.

The confirmation packets (see Figure 4.6) contain the confirmation of the receiver to the transmitter that the data connections are open or closed. Since there is a data connection for each region that needs synchronism, the connection is identified with the region identifier (ID).

Table 4.3: *Global parameters of the control data.*

Field	Content
tsize	total size of the image, width and height.
device	manufacturer and model.
delay	visualization delay.
end	end of echocardiogram transmission.

Table 4.4: *Common configuration parameters for the regions of the control data.*

Field	Content
type	type of region. Each type has a different codification, which has been described in the previous section, and needs different information. The types are: ultr (ultrasound), vid (auxiliary video), img (auxiliary image), text (text), sound (audio), ecg (ECG).
position	indicates the first coordinates of the rectangle where the region is situated.
size	indicates the size of the region, width and height.
codification	modifies the codification of the region. By default the codecs in Table 3.16.
reliability	indicates the reliability method used for that region. For instance, it may be RCVTP. The parameters of the method (if any) follow, separated by a comma.
ID	identifier of the region. This is used to identify each region and to open the UDP connections for the data of the region in the port number 1024 + ID.
mode	indicates the mode/s in which the region is visualized. The active mode is defined for the ultrasound regions. The rest of the regions can be in more than one mode. In this case, the names of the modes are separated by colons. If the region is fixed for all the modes, this field is set to 'all'.
delete	in the case that a region disappears.
mps	indicates the maximum packet size in bytes for the data packets of each region. If no value is specified no fragmentation is applied.

Table 4.5: *Region specific configuration parameters of the control data.*

Field	Content ultrasound region
color	indicates if the ultrasound has color or not.
type	indicates if the ultrasound mode is 2-D or sweep.
fps	frames per second to have temporal synchronization.
pps	swept pixels per second in the sweep modes to have temporal synchronization.
bitrate	kilobits per second used for this region.
video	to stop, go back and forward.
resolution	number of frames used together in the codification for the 2D modes or pixels per slice for the sweep modes.
calibration	how many centimeters, seconds or cm/s correspond to each pixel. The measurement is to change the units.
Field	**Content auxiliary video region**
color	indicates if the ultrasound has color or not.
fps	frames per second to have temporal synchronization.
bitrate	kilobits per second used for this region.
video	to stop, go back and forward.
resolution	number of frames used together in the codification.
Field	**Content auxiliary image region**
color	indicates if the ultrasound has color or not.
bpp	bits per pixel for this region.
codim	coded image.
Field	**Content text region**
fsize	size of the font.
Field	**Content sound region**
configuration	configuration of the codec.
channels	number of audio channels.
bitrate	kilobits per second to have temporal synchronization.
sampler	time between samples to have temporal synchronization.
Field	**Content ECG region**
cod	ECG encoding. As signal or as a sweep video.
der	number of leads.
block	size of the block in bits.
fact	factorization level for the wavelet for the 1-D codification.
sf	samples per second or pixels per second, depending on the codification, to have temporal synchronization.
time	time of ECG that is represented on the screen so as to support measurements.
bpp	bits per pixels used in the 2-D codification to have temporal synchronization.
ppslice	pixels per slice used in the 2-D codification.
width	width in centimeters so as to support measurements.

Table 4.6: *XML example for the configuration packets of the Philips Envisor device in the Figure 4.7.*


```xml
<?xml version="1.0"?>
<!DOCTYPE configuration SYSTEM "conf.dtd">


<configuration>
        <tsize w='720' h='576'/>
        <device> Philips Envisor </device>
        <region id='1' mode='Doppler', reliability='SECM,3,1,3,10,2,80,112,120'>
                <ultr color='yes' fps='25' bitrate='200' resolution='16'/>
                <position x='77' y='0'/>
                <size w='512' h='385'/>
                <mps>200</mps>
        </region>


        <region id='2' mode='B, Doppler'>
                <ecg sf='90.8' bpp='0.5' ppslice='128'/>
                <position x='494' y='52'/>
                <size w='454' h='32'/>
        </region>


        <region id='3' mode='DP', reliability='SECM,1'>
                <ultr color='no' type='sweep' fps='25' pps='121.43' bitrate='40' resolution='32'/
                    >
                <position x='200' y='0'/>
                <size w='544' h='321'/>
        </region>


        <region id='4' mode='DP'>
                <img color='no' bpp='0.5'/>
                <position x='77' y='211'/>
                <size w='150' h='130'/>
        </region>


        <region id='10' mode='all'>
                <text> PATIENT'S DATA </text>
                <position x='0' y='0'/>
        </region>
</configuration>
```

Table 4.7: *XML equivalence for real-time transmission configuration file and storage format file of the echocardiogram in Figure 3.3 captured with a Sonosite device.*

<table>
<tr><th>XML file for real-time transmission</th><th>XML files for storage</th></tr>
<tr><td><pre>
<?xml version="1.0"?>
<!DOCTYPE configuration SYSTEM "conf.dtd">
<configuration>
      <tsize w='720' h='576'/>
      <device> Sonosite </device>
      <region id='1' mode='B'>
            <ultr color='no' fps='25' bitrate
                ='200' resolution='16'/>
            <position x='76' y='52'/>
            <size w='573' h='420'/>
      </region>
</pre></td><td><pre>
<?xml version="1.0"?>
<!DOCTYPE configuration SYSTEM "roiformat.dtd">
<format>
      <tsize w='720' h='576'/>
      <region>
            <pos x0='76' y0='52' x1='659' y1
                ='472'><\pos>
            <roi><size>30083<\size><\roi>
      <\region>
<\format>
</pre></td></tr>
<tr><td><pre>
      <region id='2' mode='M'>
            <ultr color='no' type='sweep' fps
                ='25' pps='200' bitrate='40'
                resolution='32'/>
            <position x='55' y='104'/>
            <size w='512' h='456'/>
      </region>
</pre></td><td><pre>
<?xml version="1.0"?>
<!DOCTYPE configuration SYSTEM "roiformat.dtd">
<format>
      <tsize w='720' h='576'/>
      <region>
            <pos x0='55' y0='104' x1='567' y1
                ='560'><\pos>
            <roi><size>29184<\size><\roi>
      <\region>
<\format>
</pre></td></tr>
</table>

Region ID (8 bits)	Open/closed (1 bit)	...	Region ID (8 bits)	Open/closed (1 bit)

Figure 4.6: *Confirmation packets.*

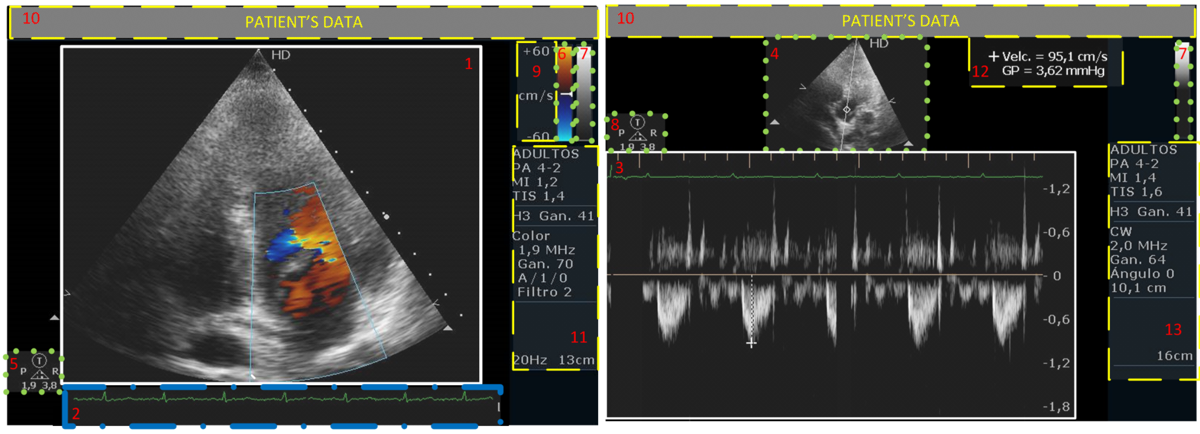


Figure 4.7: *Echocardiogram regions of a color Doppler mode (on the left) and a pulsed Doppler mode (on the right) captured with a Sonosite device. The white solid line contains the ultrasound, the blue dotted and dashed line the ECG, the green dotted line the auxiliary images and the yellow dashed line the text. Each region is identified with a number (ID) on the figure.*

4.3.3 Data Packets Coding

A UDP connexion is used for the transmission of each encoded region. However, a high data flow has to be transmitted. Furthermore, for error prone channels it is recommended to transmit small packets to enable the easy recovery of lost packets [128]. In order to support segmentation, it is necessary to establish the maximum packet size (*mps*) for the data packets and to include a sequence number in each data packet. The data packets follow the structure shown in Figure 4.8. The length of the sequence number depends on the type of region: 8 bits for sound, ECG, auxiliary video and sweep ultrasound, and 16 bits for 2-D ultrasound. The sequence number is used to distinguish each datagram of each connection that corresponds with each region. In this way, the datagrams can arrive out of order. The size of the encoded region field depends on the type of region, but it always has a size equal or inferior to the *mps* whose value is selected according to the channel and type of region. It is not necessary to specify the total size of any type of codified packets because this is known thanks to the configuration, as we will see below.

Seq no. (8 or 16 bits)	Encoded region ($\leq$ mps bits)

Figure 4.8: *Data packets.*

The coding process for each region is shown in Figure 4.9a (labelled DATA). R_T is the resolution time, referring to the time between data that is codified together. B bits are the bits after the coding. N are the packets after the fragmentation, having packets of X bits. These packets

Table 4.8: *Codification parameters expressions that correspond with the encoder in Figure 4.9.*

Parameters	Videos	Sweep videos	ECG	Sound
$\mathbf{R_T}$ **seconds**	$fps * resolution$	$\frac{resolution}{pps}$	$\frac{ppslice}{sf}$	$sampler$
B bits	$bitrate * R_T + 4$	$bitrate * R_T$	$bpp * h * ppslice$	$sampler * bitrate$
N packets	if $B \leq mps$, 1 packet (no fragmentation) if $B > mps$, $\left\lfloor \frac{B}{mps} \right\rfloor$ packets (fragmentation)			
X bits	if $B \leq mps$, $(B + header)$ bits if $B > mps$, $(N-1)$ packets with $(mps + header)$ bits and 1 packet with $(B - (N-1) * mps + header)$ bits			

correspond to the data packet shown in Figure 4.8. Thus, for each region N packets of X bits are generated every R_T seconds. It is important to highlight that when the packets are segmented, the last packet may have smaller size than the rest. These parameters, shown in Table 4.8, have different values depending on the type of region and the selected parameters in the control information (see Tables 4.3-4.5). For the 2-D modes and the auxiliary video a header of 4 bits is added to indicate how many frames are codified every R_T seconds. This header is necessary because the last time that data of a region is codified the data can contain fewer number of frames than the rest, which have $resolution$ frames. It is important to note that if some error control method is applied, the number of transmitted bits may increase and the data packets may change.

4.3.4 ETP Working Procedure

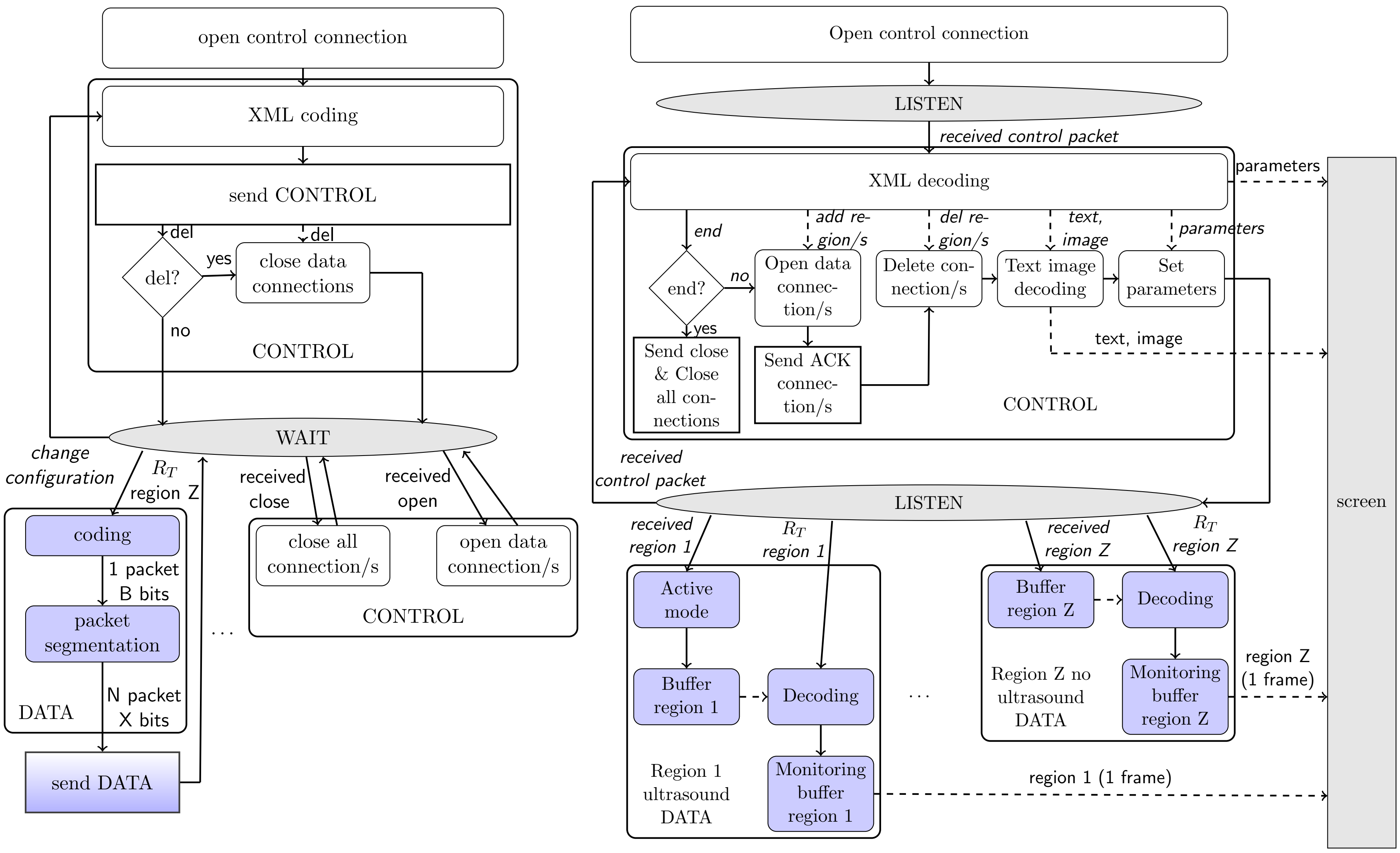
The transmitter and the receiver protocol flowcharts are shown in Figure 4.9. If an error control method is added to some region, the flowcharts will change for the respective data connections. The protocol performance for the transmitter, see Figure 4.9a, is now described.

- **Control**: the transmitter opens the control connection to send the initial configuration, where all the initial settings are defined. Then, a configuration packet is sent every time that the configuration changes, some region is added or deleted, the text or auxiliary image regions change, or the visualization ends. When a region with data connection is deleted, the data connection corresponding to that region is closed. After sending a configuration packet creating a new region with data connection, the data connection is opened after receiving the confirmation of the receiver that the data connection has been opened. After sending a configuration packet ending the transmission, all the connections are closed when the confirmation from the receiver arrives. The transmission has then finished.
- **Data**: for every region with data connection that belongs to the activated mode, the data packets are sent through its data connection each R_T seconds, except when the visualization is stopped.

The protocol performance for the receiver, see Figure 4.9b, is described below.

- **Control**: the receiver opens the control connection and listens until a confirmation packet is received. The data packets cannot be received until the first confirmation packet arrives with the initial configuration and the data connections are opened. The configuration packets can indicate the end of the transmission in which case a control packet is sent with all the closing confirmations (see Figure 4.6) and all the connections are closed. When new regions are added, the data connections are opened and a confirmation packet is sent with the opening confirmations (see Figure 4.6). If regions are deleted, the connection for those regions are closed. The text or image regions can be changed. All the parameters and new text and image regions are given to the screen in order to visualize the echocardiogram.

- **Data**: two types of data connections can be distinguished, those for ultrasound regions and those for other regions. The only difference between them is that when an ultrasound data packet is received, its mode is activated, and consequently the regions that belong to the mode are activated too. The received data packets are stored in the region buffer when a new data packet arrives. The buffer time starts the countdown when the first data packet of a mode is received. The buffer time value is sent in the configuration packets. When the buffer time expires each R_T seconds, N packets (see Table 4.8) are picked up for the activated regions to be decoded and the monitoring process begins for these regions. Each frame time the regions are picked from the monitoring buffer and are displayed on the screen.



(a) *ETP tranmitter flowchart.*

(b) *ETP receiver flowchart.*

Figure 4.9: *ETP transmitter and receiver flowchart. The dotted line represents the information flow between different parts, and the solid lines the execution procedure. The control connection is in white and the data connections in blue.*

4.4 Error Control Method for Transmission over Wireless Channels

Wireless channels are error prone, band limited and time varying. Consequently, it is necessary to introduce error control methods in order to guarantee clinical quality on reception. In [128] we proposed a Reliable Clinical Video Transmission Protocol (RCVTP). RCVTP was designed for clinical video encoded with a 3-D SPIHT algorithm, and the transmission of 2-D echocardiogram modes was tested over 3G and WiMAX channels. In this Chapter, the State Error Control Method (SECM) is proposed based on the technique described in [128] but with some added improvements. Different configurations are proposed for the data connections of ETP that correspond to different regions. The characteristics of the encoded regions are taken into account.

In [128] an error control protocol with two states that adapts to the channel conditions was proposed for the 2-D modes. The first state is used when there are few packet losses in the channel. In this state, the retransmissions mechanism is used because the use of error correction codes in these channel conditions would use a greater amount of bits than is necessary and the extra transmitted bits required would be unjustified. When more than a certain percentage of errors occur, the model turns into the second state in which both retransmissions and FEC techniques are used. The error correction code is now worth the extra amount of transmitted bits, because errors occur with the retransmissions mechanism. In this way, retransmissions adapt the bits used to the amount needed and the FEC code provides extra protection. The FEC uses Reed−Solomon (RS) code, which is a systematic block code [130].

4.4.1 SECM Working Procedure

After having explored error control methods developed over recent years, SECM is now proposed. A similar error control method to that proposed in [128] has been designed, but introducing a three states model (see Figure 4.10). Instead of using FEC and retransmissions techniques at the same time in the second state, only FEC is used. However, this has states with different FEC codes in order to adapt the bits used to the amount needed so as not to produce more delay as a result of the retransmissions. The percentage of blocks not successfully received is the number of blocks among the last 100 that have not been received successfully.

- **State 1**. This is the initial state. A FEC code is used to provide protection in case of transmission errors. However, this FEC code may not be enough. If the percentage of blocks not successfully received is more than h_{21} the model turns into *state 2*. On the other hand, if errors do not occur in the h_{10} last blocks, the bandwidth is not efficiently used and the model turns into *state 0*. This state is the initial state because it is conservative, in case errors occur.
- **State 0**. This is the state used when no errors have occurred in the last h_{10} blocks. In the event that errors start occurring and the percentage of blocks not successfully received exceeds a certain amount, the model turns into *state 1*. In this state retransmissions are used to recover the occasional errors and to use the bandwidth efficiently.
- **State 2**. This is the state when the FEC code used in *state 1* is not sufficient to recover the transmission errors and another FEC code is used with greater protection. In the event that

errors reduce in number and the percentage of blocks that are not successfully received are less than h_{21}, the model returns to the initial state, *state 1*.

More states with FEC codes having greater protection can be added to the model. The last two states can even be removed, leaving only the first state with retransmissions. The three states model is the basic proposal. The transaction parameters (see h_{01}, h_{10}, h_{12} and h_{21} in Figure 4.10 for the three states model) are selected according to the display recommendations. Two more transaction parameters are added for each added state, and two transaction parameters are removed for each state removed from the three states model. Figures 4.11 and 4.12 show the coding and decoding respectively for SECM and the three states model. N, K and *resolution* parameters are defined in Table 4.8 for ETP. Different methods are applied depending on the state. The RS code is applied to N packets that correspond to a block, so that they do not cause more coding delay and to avoid the effects of burst errors. The RS code generates K or W packets, depending on the state. N is the number of packets of the coded video and $(N-K)$ or $(N-W)$ are the number of packets of the correction. N, K or W packets are transmitted through the network, but Y packets arrive at the receiver. If N or more packets arrive correctly, the RS code is able to recover the lost packets and the video is visualized with the recommended transmission rate. Otherwise, the RS code is not able to recover the lost packets and the video is visualized with a lower transmission rate. The SECM parameters that have to be selected by the user are illustrated in Table 4.9.

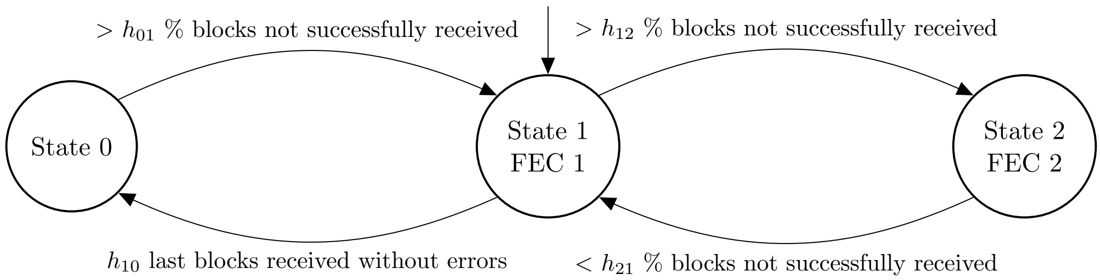


Figure 4.10: *SECM three states model.*

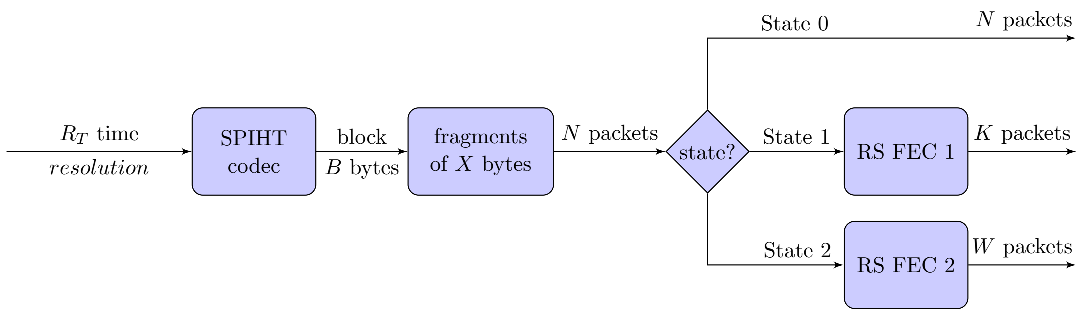


Figure 4.11: *SECM encoder for three states model.*

In order to support retransmissions, a new type of data packet has to be introduced. The acknowledge (ACK) packets (see Figure 4.13) are used to acknowledge the reception of data packets and to indicate the actual state from receiver to transmitter. The length of the sequence number for the ACK depends on the type of data packets. For ETP the length depends on the type of region: 8 bits for the sound, ECG, auxiliary video and sweep ultrasound, and 16 bits for 2-D ultrasound.

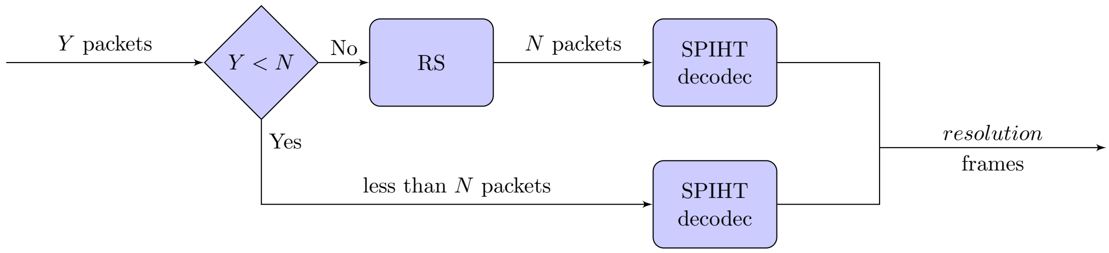


Figure 4.12: *SECM decoder for three states model.*

Table 4.9: *SECM configuration parameters.*

Parameter	Content
States	Number of states in the model. By default this value is 3.
Transactions	Transaction parameters between the states. This values correspond to h_{01}, h_{10}, h_{12} and h_{21} in Figure 4.10 for three states model. There are two parameters for each state except for the first state.
No. packets	Number of packets per resolution time that are send for each state. There are indicated as many number of packets as states in the model. This values correspond to N, K and W in Figure 4.11.

The SECM working procedure for transmitter and receiver are depicted in Figures 4.14 and 4.15. If SECM is used for the regions of ETP, a SECM transmitter and receiver have to be added for every region. The encoder and decoder are shown in Figures 4.14 and 4.15 for the three states model. The retransmission working procedure for the transmitter and the receiver, which works for *state 1*, are explained in the following paragraphs.

ACK no. 1 (8 or 16 bits)	ACK no. 2 (8 or 16 bits)	ACK no. 3 (8 or 16 bits)	ACK no. 4 (8 or 16 bits)	State (8 bits)

Figure 4.13: *ACK packets.*

- **Transmitter**. The transmitter has three possible events: each R_T second, when an ACK arrives and when retransmission time-out (RTO) expires.
  - For each frame resolution time (R_T), the ACK count is updated to the maximum number of retransmissions possible, and the data packets are generated. In order to limit the amount of transmitted bits used in channels with a high incidence of error, in which many retransmissions may be required, a maximum number of retransmissions that can be made in the frame resolution time according to the bandwidth of the network has been established. Each data packet has a sequence number that identifies it. The data packets are sent and stored in a retransmission buffer awaiting acknowledgment. A RTO timer is set each time a data packet is sent.

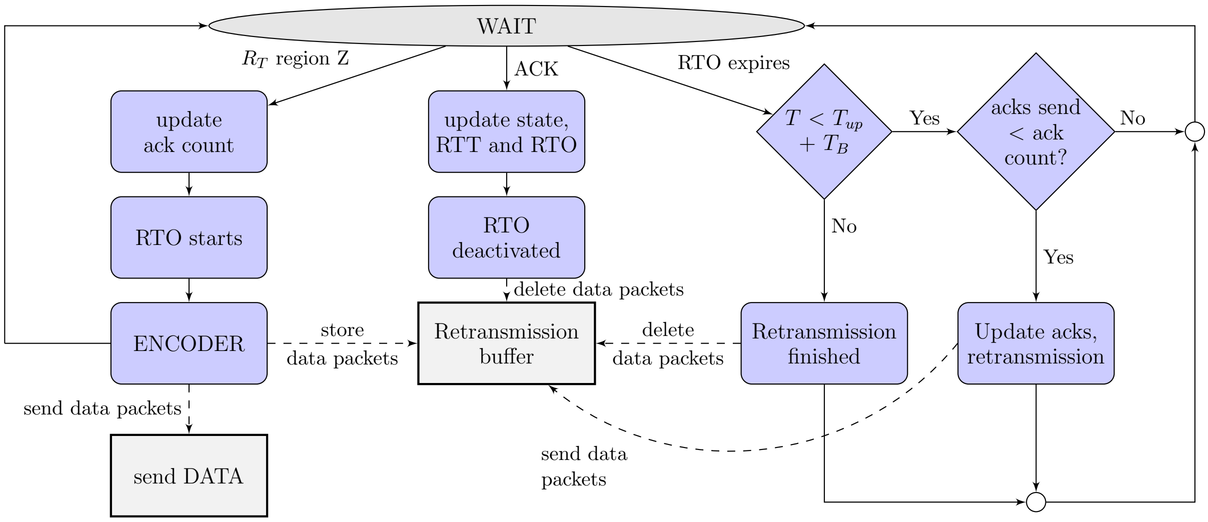


Figure 4.14: *SECM transmitter working procedure.*

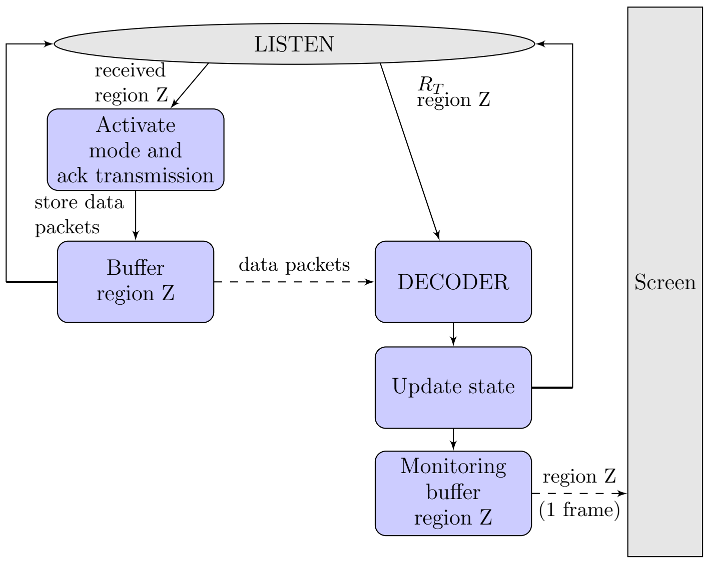


Figure 4.15: *SECM receiver working procedure.*

– Each time an ACK packet is received, the state information is picked up from the ACK packet, the RTO timers associated to the acknowledged data packet are deactivated, the data packets are removed from the retransmission buffer, and the round-trip time (RTT), transmission time coming and going, is updated. The RTO time is calculated as follows

$$RTO_i = \begin{cases} 1.1 \cdot RTT & if RTT \leq RTO_{i-1}; \\ 0.9 \cdot RTO_{i-1} & if RTT > RTO_{i-1}. \end{cases} \tag{4.2}$$

Low RTO gives low delay, but RTO values lower than RTT will lead to the duplication of packets in reception and saturate the network. Thus, the RTO value is configured to decrease each time that an ACK packet is received, but only if the RTT is higher than the RTO.

  - When an active RTO expires (no acknowledgment has been received for this data packet), the receiver checks if the current time (T) is lower than the transmission time of this packet (T_{up}) plus the buffer time (T_B) to avoid retransmitting packets that are not going to arrive in time. If the time is higher, the packet is removed from the retransmission buffer. Otherwise, the transmitter checks if the number of sent ACKs is less than the ACKs allowed, for the restriction of transmitted bits. If it is lower the retransmission is made.

- **Receiver**. The receiver has two possible events:

  - Each time a data packet is received, the current mode is updated if the region is ultrasound, an ACK packet is sent and the data packet is stored in the buffer. The buffer time starts when the first data packet is received. The monitoring process starts when the time of buffer (T_B) expires. The ACK packet acknowledges the last four data packets received and sets the current state. The mechanism of this multiple acknowledgment is similar to that used by the TCP SACK protocol [131]. Since ACK packets can also be affected by transmission errors, transmission efficiency may be improved by acknowledging not only the received packet but also some of the most recently received ones [131].
  - Each R_T time, the data packets of the next block to be visualized are picked out from the buffer. The receiver checks if all the data packets of the block have been received. If they have not, the block has not been successfully received. A frame is picked up from the monitoring buffer and is visualized.

4.4.2 SECM Configuration for ETP

Two different configurations for SECM are proposed depending on the region characteristics to be included in ETP. Each time that the echocardiogram changes the operation mode, the error control method starts working without using previous parameters, starting again from the beginning. The SECM working procedure for the receiver and transmitter (see Figures 4.14 and 4.15) has to be added to the ETP working procedure of each region (see Figure 4.9) in the part labelled DATA. One configuration is proposed for the 2-D ultrasound video and the other for the rest of the regions with data packets. The 2-D ultrasound video is the most challenging region because it contains relevant clinical information and it has higher bandwidth requirements. Both configurations are described below.

- For the 2-D ultrasound videos, a three states configuration is set. The transactions between states have been carefully chosen so as not to lose diagnostic information and use the bandwidth efficiently. The selected transaction values are $h_{01} = 1$, $h_{10} = 3$, $h_{12} = 10$ and $h_{21} = 2$.

- For the rest of the regions only one state is set. Retransmissions are used in this state because its transmission rate is very low compared to the 2-D regions, and consequently to the channel bandwidth. This leads to fewer errors and less transmission time, permitting more time for the retransmissions.

4.5 Results and Discussion for Echocardiogram Transmission

In this section the whole proposal for echocardiogram real-time transmission is tested with a WiMAX access network, using OPNET Modeler to perform the simulated transmissions.

4.5.1 Evaluation Setup for Echocardiogram Transmission

The echocardiogram database described in Section 3.1.2.1 has been used for transmission in the simulation scenarios. Important aspects of the database are the echocardiogram time and video distribution, each video corresponding to a change of operation mode (see Table 3.4), and the regions for each type of device and visualization mode (see Table 3.5). Other important aspects are the quality parameters to be evaluated. In this section, the simulation scenarios, the quality parameters and the ETP and SECM selected parameters are described.

4.5.1.1 Simulation scenarios

The two simulation scenarios are illustrated in Figure 4.16. The first scenario has fixed access and the second has mobile access, both to a WiMAX network. The first scenario is used to select the ETP and SECM parameters and the second scenario to test the proposed methods working separately and as a whole. The echocardiogram acquisition is thus performed in a remote location with fixed or mobile WiMAX access connected to the Internet. The echocardiogram is sent in real-time to the expert cardiologist located in a hospital with an Asymmetric Digital Subscriber Line (ADSL) access where the diagnosis is made.

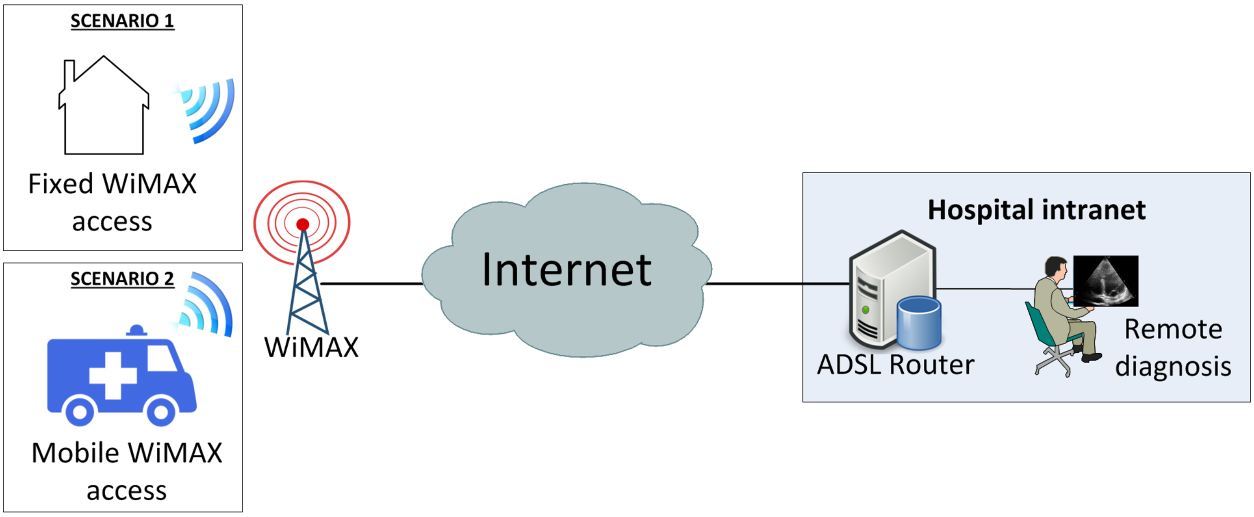


Figure 4.16: *Simulation scenarios for tele-echocardiography with WiMAX access.*

The scenarios have been created using OPNET modeler. The WiMAX (IEEE 802.16e) configuration parameters are presented in Table 4.10 and the WiMAX QoS parameters in Table 4.11. The distance between the access point and the WiMAX base station for the fixed scenario is 1 km. For the mobile scenario the distances are from 100 meters up to 2 km. The speed of the vehicle for the mobile scenario ranged between 60 km/h and 120 km/h. This configuration covers the possible mobility cases and a wide variety of channel conditions. The WiMAX base station is connected

Table 4.10: *WiMAX configuration parameters.*

Parameter	Value
Access technology	OFDMA 5 MHz
Base frequency	2.5 GHz
Frame/symbol duration	5 ms/100.8 μs
Modulation and coding	QPSK 1/2
Duplexing technology	TDD
Multipath channel model	ITU vehicular B
Pathloss model	Vehicular environment

Table 4.11: *WiMAX QoS parameters.*

Parameter	Value
Maximum sustained traffic rate	1 Mbps
Maximum reserved traffic rate	0.5 Mbps
Maximum Latency	30 ms

to the Internet via a Digital Signal 3 (DS3) Wide Area Network (WAN) link (44.736 Mbps). The hospital intranet has a router that connects to the Internet cloud through a DS3 link. The hospital intranet has a bandwidth of 100 Mbps. The approximate distance between the two subnets is 3342 km, which corresponds to approximately 11.1 ms propagation delay. The packet loss ratio in the Internet cloud is 0.001 %. The cloud has a 1 ms delay in addition to the propagation delay of the WAN link.

4.5.1.2 Quality Parameters

The following parameters measure the quality of the transmission:

- *Transmitted bandwidth (BW_{transm}):* quantifies the number of bits per second (bps) used in the application layer of the communication. It is used to measure the efficiency of the protocol in the transmission process. Note that retransmissions and FEC techniques increase the transmitted bandwidth.

- *Effective bandwidth ($BW_{effective}$):* quantifies the useful number of bits per second (bps) used to decode the ultrasound region. Note that the effective bandwidth does not correspond to the transmitted bandwidth due to errors in the transmission, retransmissions and FEC. As the ultrasound regions are compressed with SPIHT algorithm and it is embeddedness, when an error occurs the rest of bits are ignored and the region is decompressed with the received previous bits. This parameter is used to measure the quality of the display echocardiogram, and to check if the display recommendations are achieved for each fragment that corresponds to a mode.

- *Percentage of time with guaranteed clinical quality:* This is the percentage of the time that the ultrasound region of the echocardiogram is visualized with the display recommendations.

- *Number of fragments without guaranteed clinical quality:* This is the number of fragments that do not fulfill the display recommendations. Each fragment is checked separately. The echocardiograms are composed of fragments that correspond to a change of operation mode (number of videos in Table 3.4).

- *Delay:* This is the time from the moment when the video is captured until it is visualized in the receiver. A simplified point-to-point delay equation would include the following terms

$$T_D = T_B + T_{1p} + R_T + T_{pr} \tag{4.3}$$

  where T_D is the delay time, T_B is the buffer time, T_{1p} is the transmission time of the first packet (this depends on the network and fragment size), R_T is the resolution time, and T_{pr} is the processing time. The longest times are the R_T, which is 0.64 s for the 2-D modes, and T_B, which is determined by the user.

4.5.1.3 Parameters Selection

The ETP configurations for the main regions are shown in Table 4.6 for the Philips Envisor and in Table 4.7 for the Sonosite device. The encoding parameters are those recommended in Chapter 2 (see Table 3.16). The main ETP parameters were described in Table 4.8 and the SECM parameters in Table 4.9. There are some ETP and SECM configuration parameters, such as the maximum packet size, the FEC codes and the buffer time, that are very important for the good performance of the echocardiogram transmission. In order to choose suitable parameters, some simulations have been performed in the fixed scenario described previously. These simulations are performed for a 2-D ultrasound video of 1 minute of duration, since this region presents the most demanding characteristics. The error control methods used are various FEC codes since FEC is the core technique of SECM. The simulations have been carried out for maximum packet sizes (*mps*) of 200, 400 and 800 bytes, FEC codes of 40% and 50% of protection (FEC 1 and FEC 2, respectively), and buffer times (T_B) of 0.4, 0.45 and 0.5 seconds. Whether the diagnosis with the mode is possible or not is shown in Table 4.12 for the two FEC codes. The transmitted bandwidth is shown in Table 4.13, this being independent of the buffer time.

Table 4.12: *Is diagnosis possible for a 2-D ultrasound mode region with the fixed scenario and correction codes of 40% and 50 %?*

$\mathbf{T_B}$	mps (bytes) for FEC 1			mps (bytes) for FEC 2		
	200	**400**	**800**	**200**	**400**	**800**
0.4	NO	NO	NO	NO	NO	NO
0.45	YES	NO	NO	YES	YES	NO
0.5	YES	YES	NO	YES	YES	NO

The selected ETP and SECM parameters for the different regions of the available devices are shown in Table 4.14 and discussed below.

Table 4.13: *Transmitted bandwidth in kbps for a 2-D ultrasound mode region with the fixed scenario.*

FEC code	mps (bytes)		
	200	**400**	**800**
FEC 1: 40%	281	279	278
FEC 2: 50%	301	299	298

- **Maximum packet size (*mps*).** The encoded blocks are fragmented in packets of *mps* bytes. The *mps* value changes for each region. Fragmentation is applied for the 2-D mode ultrasound video regions only. Since the rest of the regions have a block size lower than 1500 bytes and these regions are less affected by errors than the 2-D regions, fragmentation is not necessary. The selected *mps* value for the 2-D modes ultrasound video is of 200 bytes, as in [128]. This value was selected because although it introduces more overhead than higher values (see Table 4.13), this value is less affected by errors and the errors are more easily recovered (see Table 4.12). This leads to a value of 80 packets for N.

- T_B **seconds**. A buffer time of 0.45 seconds is selected. This is the lowest value that allows the errors introduced by the channel to be recovered (see Table 4.12). The same T_B value has to be selected for all the regions in order to display them all with the same delay.

- **FEC codes**. FEC codes are used for the 2-D ultrasound regions only. The selection of protection codes is restricted as the available codes depend on the fragment size (*mps*). For the packet size used in this Thesis, 200 bytes, the available protection codes are 40 %, 50 % and 75% of the total block size. Thus we have chosen 40 % for *FEC 1* and 50 % for *FEC 2*. Both correction codes introduce enough protection to guarantee clinical quality (see Table 4.12) for the simulated scenario, which corresponds to 11 % of lost packets. These parameters lead to values of 112 packets for K and of 120 packets for W.

4.5.2 Results for Echocardiogram Transmission

In this section the proposed system as a whole is tested with the evaluation setup described in the previous section for echocardiogram transmission. The different proposals for each part are tested separately in order to see how each proposed part contributes to the improvement of the overall system. Furthermore, several SECM configurations are compared in order to verify the improvements of the proposed configuration. The results have been split into two sections, the first without using SECM and the second using SECM. In both sections the results using ETP and without using ETP are also shown.

4.5.2.1 Results for Echocardiogram Transmission without SECM

No error control method is used in this section to test the advantages of using ETP, and consequently of using different codification and transmission techniques for each type of region.

Table 4.14: *ETP and SECM parameters for the different regions. S., P. Envisor and P. IE33 are the three available devices.*

Parameters	2-D ultrasound	Sweep ultrasound (S./P. Envisor/P. IE33)	ECG (S./P. Envisor/P. IE33)
R_T (s)	0.64	0.16/0.28/0.18	-/1.69/1
mps (bytes)	200	-	-
B (bytes)	2000	800/1400/900	-/512/656
X (bytes)	200	800/1400/900	-/512/656
N (packets)	80	-	-
K (packets)	112	-	-
W (packets)	120	-	-
T_B (s)	0.45	0.45	0.45

The nine available echocardiograms have been transmitted using the mobile scenario under two different configurations:

- **Without ETP.** The echocardiograms have been transmitted without distinguishing modes or regions. The echocardiogram is sent using the same codification algorithm, transmission rate (200 kbps) and configuration as for the 2-D ultrasound regions in ETP, since these regions are the most restrictive. A sequence number of 16 bits is added to the encoded part for each sent packet as for 2-D ultrasound regions in ETP to allow out of sequence arrival of the packets.

- **With ETP.** The echocardiograms have been transmitted distinguishing modes and regions, using ETP but without using an error correction method.

In Tables 4.15, 4.16 and 4.17 the transmitted bandwidth, the percentage of time with guaranteed clinical quality and the number of fragments without guaranteed clinical quality are shown for both configurations. As we can see in Table 4.15, using ETP, less bandwidth is transmitted than without using ETP for all the available echocardiograms. The transmitted bandwidth without considering modes is the same for all the echocardiograms, 200 kbps of encoded video plus the sequence number header. However, the transmitted bandwidth when ETP is used depends on the mode distribution of the echocardiograms. The control packets used in ETP represent less than 4 % of the total transmitted data for the nine echocardiograms. In general, the higher the percentage of time that the 2-D modes are presented in the echocardiogram, the higher is the transmitted bandwidth. The 2-D ultrasound regions have a transmission rate of 200 kbps, while the sweep ultrasound regions have a transmission rate of 40 kbps. The bandwidth saving using ETP ranges from 15 kbps (echocasrdiogram 9) to 81 kbps (echocasrdiogram 7) for the available echocardiograms. As we can observe in Table 4.16, the percentage of time with guaranteed clinical quality using ETP is higher than without using ETP for all the available echocardiograms. The increase in the percentage of time ranges from 6 % (echocardiogram 9) to 19 % (echocardiogram 8). As we can see in Table

4.17, no fragments are visualized with guaranteed clinical quality for the configuration without distinguishing modes. The number of fragments without guaranteed clinical quality are the total number of fragments for each echocardiogram (see number of videos in Table 3.5). The number of fragments without guaranteed clinical quality using ETP has dropped by at least 10 fragments, 53 % of the total fragments.

Although the echocardiograms are visualized with guaranteed clinical quality for higher percentages of time when ETP is used than without using ETP, between 16 % and 25 % of the time the available echocardiograms are visualized without guaranteed clinical quality using ETP. Furthermore, between 2 and 9 echocardiogram fragments are not visualized with guaranteed clinical quality using ETP.

Table 4.15: *Transmitted bandwidth without using SECM.*

Configuration	Bandwidth per echocardiogram (kbps)								
	1	2	3	4	5	6	7	8	9
without ETP	202	202	202	202	202	202	202	202	202
with ETP	147	123	139	159	174	159	121	171	187

Table 4.16: *Percentage of time with guaranteed clinical quality without using SECM.*

Configuration	Percentage of time per echocardiogram								
	1	2	3	4	5	6	7	8	9
without ETP	76.58	75.76	83.61	77.13	74.92	75.91	77.64	74.65	74.54
with ETP	90.47	82.94	93.22	87.73	92.2	88.52	94.21	93.77	80.7

Table 4.17: *Number of fragments without guaranteed clinical quality without using SECM.*

Configuration	No. of fragments per echocardiogram								
	1	2	3	4	5	6	7	8	9
without ETP	18	20	19	25	22	25	25	41	20
with ETP	5	6	9	3	5	8	5	2	9

In Figures 4.17, 4.18 and 4.19 the effective bandwidth over the time is shown without using ETP for echocardiograms numbers 3, 7 and 9, respectively. The effective bandwidth when the fragments are not visualized with sufficient clinical quality is shown in red and when the fragments are visualized with sufficient clinical quality is shown in blue. The expected effective bandwidth for each fragment in order to visualize the echocardiogram with sufficient clinical quality is shown in gray. We can observe in these figures that the case without using ETP (see Figure 4.17) is the most affected by errors since modes are not distinguished and the highest bandwidth is transmitted independently of the operation mode. In the case where the operation modes are distinguished, (see Figures 4.18 and 4.19) the 2-D modes, the modes with the highest transmission rate, are more

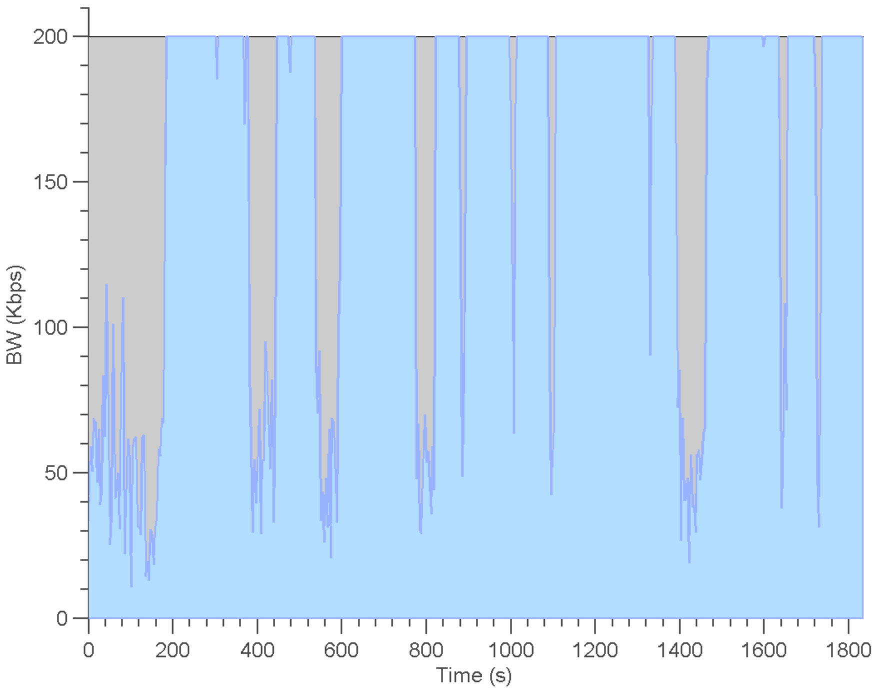


Figure 4.17: *Effective bandwidth over the time without using SECM or ETP for echocardiogram number 3 is drawn in blue and the expected effective bandwidth to visualize the echocardiogram with sufficient clinical quality is drawn in gray.*

affected by the errors of the channel than the other modes with lower transmission rates. Note that there are fragments that are visualized with sufficient clinical quality although the effective bandwidth is lower than the expected bandwidth due to display recommendations.

4.5.2.2 Results for Echocardiogram Transmission with SECM

Different error control methods for the 2-D ultrasound region are compared in this section in order to identify which is the most suitable. Different configurations and modifications to the basic three states SECM have been tested, using one-state, two-state and three-state models. Furthermore, FEC and retransmission techniques have been tested in the different states and even working both techniques together. When ETP is used, the error control method for the sweep ultrasound regions is SECM with one state model with retransmissions. The nine available echocardiograms have been transmitted using the mobile scenario, with and without using ETP and using the following error control methods:

- **One state**:
  - FEC 1 (40% of protection).
  - FEC 2 (50% of protection).
- **Two states**:
  - *State 1*: ACK, *state 2*: FEC 1 (40% of protection).

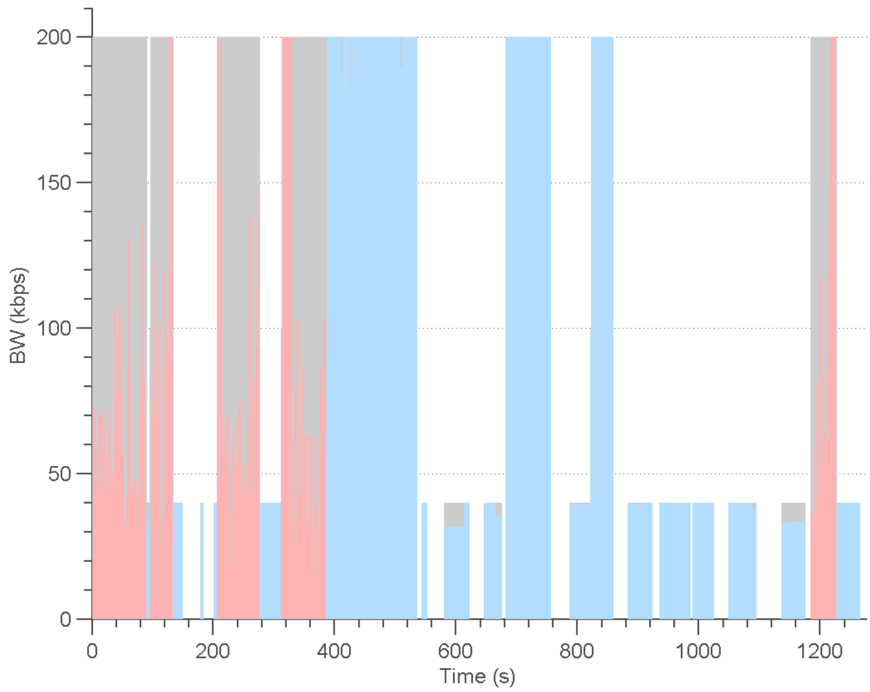


Figure 4.18: *Effective bandwidth over the time without using SECM and using ETP for echocardiogram number 7. The effective bandwidth when the fragments are not visualized with sufficient clinical quality is shown in red and when the fragments are visualized with sufficient clinical quality is shown in blue. The expected effective bandwidth for each fragment in order to visualize the echocardiogram with sufficient clinical quality is shown in gray.*

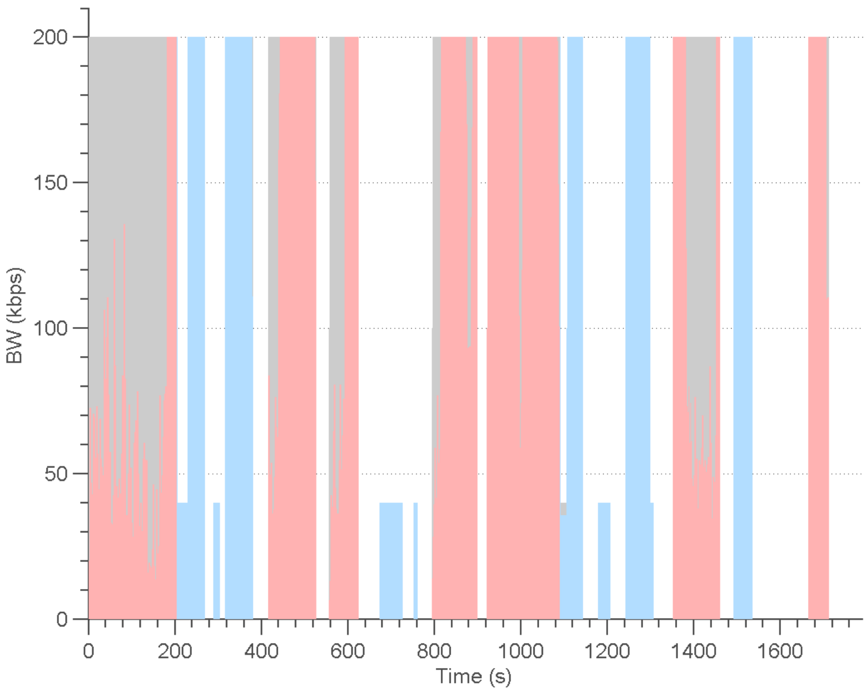


Figure 4.19: *Effective bandwidth over the time without using SECM and using ETP for echocardiogram number 9. The effective bandwidth when the fragments are not visualized with sufficient clinical quality is shown in red and when the fragments are visualized with sufficient clinical quality is shown in blue. The expected effective bandwidth for each fragment in order to visualize the echocardiogram with sufficient clinical quality is shown in gray.*

  - *State 1*: ACK, *state 2*: FEC 2 (50% of protection).
  - *State 1*: ACK, *state 2*: ACK + FEC 1 (40% of protection).
  - *State 1*: ACK, *state 2*: ACK + FEC 2 (50% of protection).

- **Three states**:

  - *State 1*: ACK, *state 2*: FEC 1 (40% of protection), *state 3*: FEC 2 (50% of protection).

Tables 4.18 and 4.19 show the transmitted bandwidth and the number of fragments without guaranteed clinical quality for the different error control methods and the nine available echocardiograms using ETP. When no fragments are visualized without guaranteed clinical quality, then 100 % of the time the echocardiogram is visualized with guaranteed clinical quality. As we can see, the proposed error control method with three states uses the lowest bandwidth displaying all the fragments with adequate clinical quality for the nine echocardiograms. All the fragments are correctly visualized with only one state and both FEC codes. The drawback of using only one state is that more bandwidth than necessary is used when the channel introduces few errors. In the case of the two states model, the only configuration which achieves all the fragments with adequate clinical quality is that using the most protective FEC code. When a less protective FEC code is used, less bandwidth is transmitted, but not all the fragments are visualized correctly for the nine echocardiograms. When retransmissions and the FEC techniques are used at the same time in the second state, as in [128], more bandwidth is used and more fragments without clinical quality are displayed than when using the FEC technique alone, because a longer buffer time is necessary to recover the errors with the retransmissions. The best results are obtained using SECM and the proposed three states model because the state model adapts to the channel conditions. All the fragments are visualized with adequate clinical quality and the transmitted bandwidth ranges from 154 kbps to 244 kbps for the available echocardiograms.

Table 4.18: *Transmitted bandwidth for various error control methods and ETP.*

Methods		Bandwidth per echocardiogram (kbps)								
		1	2	3	4	5	6	7	8	9
1 state	**FEC 1**	203	168	191	222	241	221	163	238	260
	FEC 2	216	179	204	237	258	236	173	255	278
2 states	**ACK, FEC 1**	187	156	173	215	222	204	153	221	240
	ACK, FEC 2	197	162	181	228	232	213	162	233	251
	ACK, ACK+FEC 1	201	165	181	232	235	215	162	234	255
	ACK, ACK+FEC 2	207	170	187	245	244	225	170	244	265
3 states	**ACK, FEC 1, FEC 2**	189	157	174	225	226	206	154	222	244

Figures 4.20 and 4.21 show the transmitted and effective bandwidth over the time for the proposed transmission method using ETP and SECM with three states for two different echocardiograms. If we compare these figures with Figures 4.18 and 4.19 we can see that more bandwidth is used in the modes with more errors, where the highest bandwidth is used due to the FEC code. Therefore, the three states model adapts to the channel conditions.

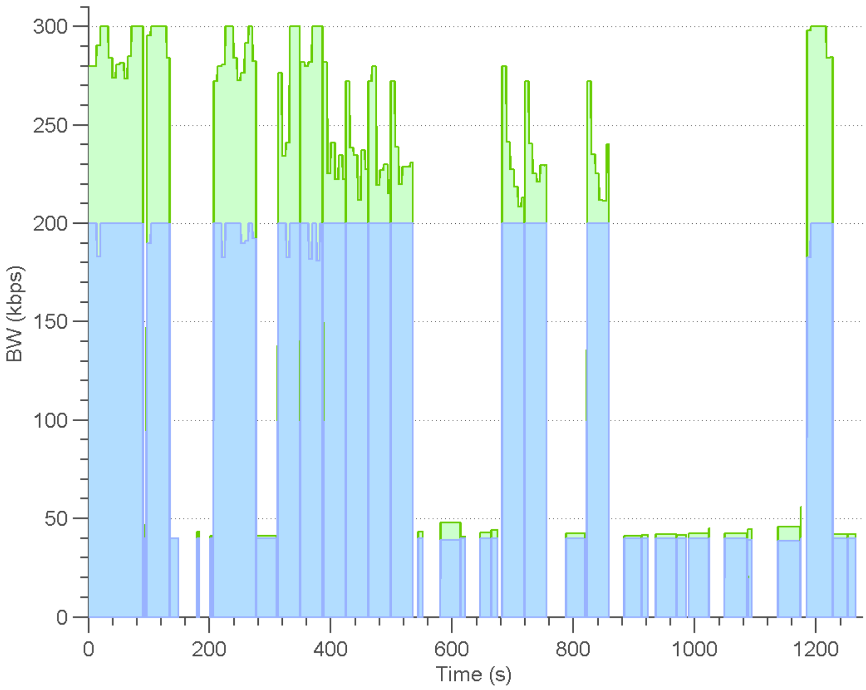


Figure 4.20: *Effective and transmitted bandwidth over the time with ETP and SECM for echocardiogram number 7. The effective bandwidth is shown in blue and the transmitted bandwidth in green.*

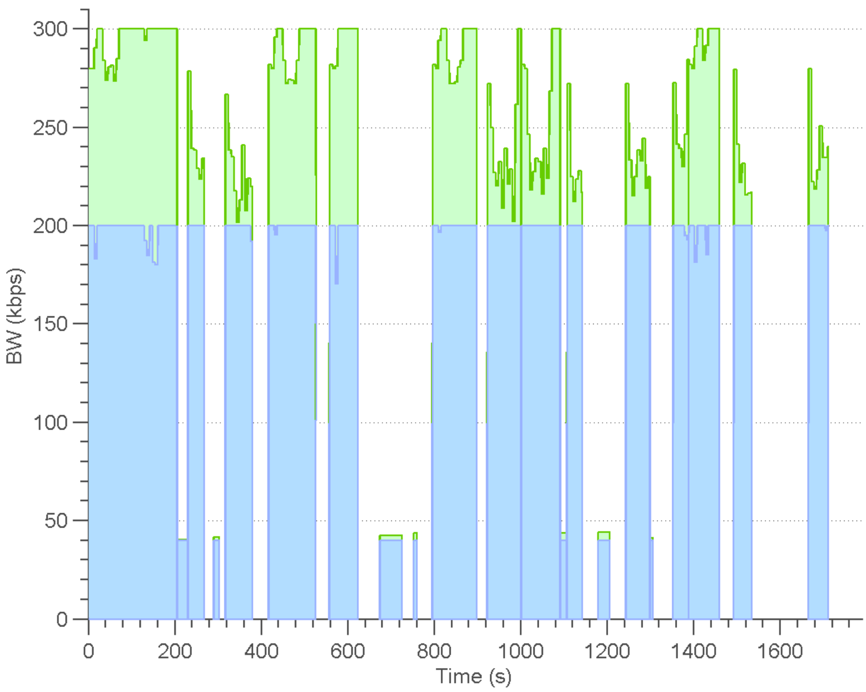


Figure 4.21: *Effective and transmitted bandwidth over the time with ETP and SECM for echocardiogram number 9. The effective bandwidth is shown in blue and the transmitted bandwidth in green.*

Table 4.19: *Number of fragments without guaranteed clinical quality for various error control methods and ETP.*

Methods		Number of fragments per video								
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1 state	**FEC 1**	0	0	0	0	0	0	0	0	0
	FEC 2	0	0	0	0	0	0	0	0	0
2 states	**ACK, FEC 1**	0	0	0	0	0	0	0	1	0
	ACK, FEC 2	0	0	0	0	0	0	0	0	0
	ACK, ACK+FEC 1	0	0	1	0	0	0	0	1	0
	ACK, ACK+FEC 2	0	0	0	0	0	2	0	1	0
3 states	**ACK, FEC 1, FEC 2**	0	0	0	0	0	0	0	0	0

Table 4.20 shows the transmitted bandwidth and the maximum number of fragments without guaranteed clinical quality of the nine echocardiograms for various error control methods when ETP is not used (all the echocardiograms are considered as 2-D mode). The transmitted bandwidth is the same for all the echocardiograms since it is independent of the operation mode distribution without using ETP. Without using ETP more bandwidth is transmitted than using ETP, and consequently the echocardiogram visualization is more affected by the errors. The only error control methods that guarantee adequate clinical quality for the whole echocardiogram is when FEC 2 is used and with the 3 states model. For these two methods the bandwidth used is similar because for the three states method the model is in state 3 almost all of the time when the FEC 2 code is applied. The transmitted bandwidth for the three states model is 308 kbps for all the echocardiograms.

Table 4.20: *Transmitted bandwidth and maximum number of fragments without guaranteed clinical quality for various error control methods and without using ETP.*

Methods		Bandwidth	# fragments
1 state	**FEC 1**	286	3
	FEC 2	310	0
2 states	**ACK, FEC 1**	287	3
	ACK, FEC 2	307	2
	ACK, ACK+FEC 1	313	2
	ACK, ACK+FEC 2	335	2
3 states	**ACK, FEC 1, FEC 2**	308	0

4.5.3 Discussion for Echocardiogram Transmission

The use of ETP allows a saving in the transmitted bandwidth, as is shown in Table 4.15, since the visualization characteristics are taken into account in the compression and transmission of the echocardiograms. The saving of bandwidth depends on the mode distribution of the echocardiograms. The 2-D modes present the highest transmission rate while the sweep modes present a very low transmission rate. Hence the fact that a higher percentage of 2-D modes leads to higher transmitted bandwidth. The modes with lower transmission rates are less affected by the errors introduced by the channel, as is shown in Figures 4.17, 4.18 and 4.19. Therefore, the lower the transmitted bandwidth, the higher is the percentage of time with guaranteed clinical quality, as shown in Tables 4.15 and 4.16. If the regions and modes are not taken into account, the highest bandwidth has to be used for all the modes, the echocardiograms being more affected by the transmission errors. The percentage of time with adequate clinical quality using ETP increases between 6 % and 19 %, while the number of fragments without adequate clinical quality using ETP decreases between 10 and 39 compared to not using ETP, as is shown in Tables 4.16 and 4.17, respectively. The percentage of time with guaranteed clinical quality depends on the error distribution and the time distribution of the operation modes, especially on the time distribution of the 2-D modes. The number of fragments that are visualized without guaranteed clinical quality not only depends on the error distribution but also on the number of fragments that correspond with 2-D modes.

Although the echocardiograms are less affected by errors when ETP is used, there are still some parts that are visualized without guaranteed clinical quality when the channel introduces errors. Therefore, an error control method is necessary in order to guarantee clinical quality when transmission errors occur. SECM has been designed to adapt the error control method to the channel conditions with the proposed configuration. Error control methods increase the transmitted bandwidth in order to protect against errors. If FEC is used in a channel without errors, the bandwidth is inefficiently used. Retransmissions can cause long delay when many errors occur. Thus, in the proposed state model the retransmission method is used when few errors occur. When retransmissions are not sufficient, the FEC method is used with different correction codes. The selected configuration for SECM has been demonstrated to adapt to the channel conditions, as is shown in Figures 4.20 and 4.21. Furthermore, if the proposed SECM configuration is used, the echocardiograms are visualized with adequate clinical quality for a mobile network with WiMAX access and representative settings, as can be seen in Tables 4.19 and 4.20. However, without SECM, the available echocardiograms are visualized without guaranteed clinical quality for between 16.39 % and 25.46 % of the time, as shown in Table 4.16.

If both techniques ETP and SECM are used together, a saving in the transmitted bandwidth is achieved compared to not using ETP. Furthermore, all the available echocardiograms are received with guaranteed clinical quality with the proposed SECM configuration. The bandwidth saving depends on the modes distribution of the echocardiogram and on the error distribution. For a mobile network with WiMAX access and representative settings, the transmitted bandwidth is between 154 kbps and 244 kbps for the available echocardiograms , as shown in Table 4.18. However, if ETP is not used and the echocardiogram is transmitted without distinguishing modes, the transmitted bandwidth is 308 kbps, as can be seen in Table 4.20. Thus there is a bandwidth saving of up to 154 kbps, 100 % of bandwidth saving. Without ETP the operation modes are not distinguished, all the modes are considered as 2-D modes and consequently more bandwidth is transmitted. Furthermore, blocks are not successfully received due to the fact that 2-D modes are more affected by errors and that the model for all the echocardiograms is in the last state (with more protection) almost all the time, which transmits more packets.

A comparison of the results with previous works (see Table 2.3) shows that the bandwidths using a WiMAX channel in the previous works are higher than that used with the proposed system, even for clinical videos with lower resolution. Comparing the results for clinical videos with a resolution similar to the echocardiograms used in this Thesis and having acceptable clinical quality on reception reveals that the saving in the bandwidth used in the present work is greater than 1 Mbps. The saving is due to several factors:

- The compression method based on regions takes into account the visualization characteristics of the echocardiograms, hence the echocardiogram is compressed efficiently.
- The compression recommendations give the minimal transmission rate to guarantee clinical quality for the ultrasound regions of each operation mode.
- ETP allows the transmission by regions, and consequently compresses each region with its minimal recommended bandwidth.
- SECM adapts the error control method to the channel conditions using the bandwidth more efficiently.
- The display recommendations establish whether the echocardiogram is visualized with clinical quality even if errors occur, and as a result not all the errors have to be recovered to be able to visualize the echocardiogram with guaranteed clinical quality.

4.6 Conclusions

The main objective of the work described in this Chapter is to achieve the transmission and the display of the echocardiograms without losing diagnostic information. In order to fulfill this objective, three tasks have been carried out: the design of a clinical evaluation for the display recommendations, a proposal of a protocol for real-time transmission of echocardiograms compressed by regions, and the design of an error control method. The main conclusions resulting from these three tasks are the following:

- An evaluation methodology has been designed to give recommendations for the display of echocardiograms after transmission with errors. The methodology has been specially designed for the echocardiogram compression methodology proposed in this Thesis based on visualization modes. However, any medical video can be evaluated using the evaluation designed for the 2-D modes. The proposed evaluation is easy to perform because it starts from the previous recommendations given for the transmission rate for compression. The evaluation gives the display recommendations, the time and the range of the transmission rates for the visualization of the video with a lower transmission rate than the recommended rate due to transmission errors. It has a major advantage over making an evaluation for each visualized echocardiogram after transmission. With only one evaluation we are able to know if the echocardiogram is visualized with adequate quality without the need to carry out an assessment for each transmission. Display recommendations are given for the echocardiogram compressed with the proposed method. For the 2-D modes, the ultrasound part can be visualized at any transmission rate up to 5 % of the time. For the sweep modes, up to two image slices (about 10 % of the screen) can be visualized at any transmission rate.

- A protocol for the end-to-end transmission of echocardiograms in real-time has been designed, known as the Echocardiogram Transmission Protocol (ETP). ETP allows transmitting the echocardiogram compressed by regions. The regions are transmitted separately, and consequently different transmission rates and error control methods can be used depending on the clinical importance of the region and on the network. ETP can be used for transmission in any network. The simulated transmissions have demonstrated that by transmitting the echocardiogram by regions, less bandwidth is used and the echocardiogram is visualized with adequate clinical quality for a greater percentage of time than when transmitting without considering the regions and modes.

- An error control method based on states is proposed, known as the States Error Control Method (SECM). SECM adapts to the channel conditions using different states depending on the errors occurring. Furthermore, different numbers of states can be set depending on the data and the network characteristics. SECM has been demonstrated to adapt the bandwidth used to the channel conditions and to guarantee quality on reception. This method can be used for any data to be transmitted over error prone channels.

Chapter 5

Conclusions and Future Work

This last Chapter sets out the conclusions of the Thesis and future work. The research objectives presented in Chapter 1 have been discussed throughout the Thesis. The most challenging key factors in tele-echocardiography systems have been addressed. These include encoding algorithms, transmission protocols, error control methods, wireless transmission technologies and clinical quality. This Chapter is organized as follows. Section 5.1 describes the objectives achieved in the Thesis chapter by chapter. Section 5.2 enumerates the contributions of this work and the accomplished results. Finally, future work is described in Section 5.3.

5.1 Research Objectives Achieved

Chapter 1 describes the motivations and evidence of the benefits that telemedicine systems in general and tele-echocardiograhy systems in particular can provide to patients and to the healthcare system. The main challenges and critical issues involved are described in Section 1.2 and the objectives are established in Section 1.3. The main aim of this Thesis is to investigate telemedicine systems applied in cardiology environments. All the detailed objectives listed in Chapter 1 are addressed throughout the remaining chapters.

- Reviews on the state of the art in general aspects of compression methods for both image and video, clinical quality, wireless technologies and transmission protocols have been carried out in Chapter 2. Specifically, a review of the literature on compression and transmission techniques used in wireless medical video transmission systems is also described in Chapter 2.
- The design, evaluation and recommendations for use of compression methods for the storage and real-time transmission of echocardiograms are addressed in Chapter 3.
- The design and evaluation of protocols for the transmission of echocardiograms in real-time and recommendations for the echocardiogram visualization after transmission are addressed in Chapter 4.

The detailed conclusions and objectives achieved of each section and chapter are listed in Sections 2.8, 3.7 and 4.6.

5.2 Contributions and accomplished results

The major contribution of this Thesis is to provide support to tele-echocardiography systems. The most challenging issues have been addressed and efficient solutions have been proposed. This Thesis thus provides a framework for the real-time transmission and storage of echocardiograms, preserving diagnostic information and making an efficient use of resources (disk space and transmission rate). The work involved in achieving this main objective has also resulted in several minor contributions:

- A two phase evaluation methodology for the compression and transmission of clinical images has been designed. This methodology is accurate but less burdensome than other tests proposed in the literature. The evaluation consists of two tests. The first provides compression recommendations and the second provides display recommendations. These tests enable the transmission of clinical videos with a minimal transmission rate and establish whether the clinical image is visualized without losing diagnostic information and without the need to carry out other evaluations for each transmitted video. The tests have been designed for an echocardiogram compression method based on visualization modes, but the same methodology can be used for other clinical images and compression methods. The testbed for the compression recommendations has been split into two, maintaining the degree of accuracy but reducing the assessment process. The starting point of the evaluation is the recommendations given for the compression in order to simplify the assessment process.

- An image compression format for storage purposes has been proposed that saves storage space with respect to the conventional image formats incorporated in the DICOM standard. The proposed storage format takes advantage of the segmentation facilities of the device to enhance the compression performance without adding complexity to the device that already incorporates the DICOM standard.

- A tool that makes conversions between different image formats has been developed. This tool allows interoperability between medical centers and devices, and also enables echocardiograms to be stored with the proposed format and thus saves storage space. Options to edit the image (edit/add/remove regions) taking advantage of the proposed format based on regions have been also included.

- An echocardiogram compression method for real-time transmission purposes based on regions and visualization modes has been designed. An encoded algorithm is proposed according to the type of data and visualization characteristics for each region. This compression method that takes advantage of the segmentation facilities and the visualization characteristics of the devices has demonstrated that it compresses echocardiograms more efficiently than conventional video codecs. Furthermore, each region can be compressed with different transmission rate according to its clinical quality. Very good results in terms of bandwidth use have been obtained, especially for the M and pulsed/continuous Doppler modes, thanks to the fact that the compression method takes into account the stationary characteristics of the sweep modes and only a thin slice is compressed for every frame.

- An Echocardiogram Transmission Protocol (ETP) for end-to-end and real-time transmission of echocardiograms compressed by regions has been designed. The regions are transmitted separately, and consequently different transmission rates and error control methods can be used depending on the clinical importance of the region and on the network. Therefore, ETP can be used for transmission in any network. The simulated transmissions have demonstrated that by transmitting the echocardiogram by regions, less bandwidth is used and the echocardiogram is visualized with clinical quality for a greater percentage of time when errors occur than without considering the regions and modes.

- A States Error Control Method (SECM) that adapts to the channel conditions has been proposed. Different error control methods are used depending on the channel errors. Furthermore, different configurations can be set depending on the data and the network characteristics. SECM has been demonstrated to adapt the transmitted bandwidth to the channel conditions and to guarantee quality on reception. This method can be used for any type of data to be transmitted over error prone channels.

- An overall system for real-time echocardiogram transmission over wireless networks has been proposed. The system makes use of all the previous proposals, compression, ETP, SECM and clinical evaluation. The configuration parameters for the echocardiogram transmission in a network with WiMAX access have been proposed. The proposed system and configurations have been demonstrated to save transmitted bandwidth with the echocardiogram being received with guaranteed clinical quality when compared to other compression techniques that do not distinguish regions and visualization modes, and compared to other error control methods.

- A general methodology to design a framework for the storage and transmission of any clinical video has been proposed. This methodology guarantees the reception of the video without losing diagnostic information and making efficient use of resources as a result of taking into account the data characteristics to compress the video efficiently, carrying out an evaluation, and designing transmission techniques that adapt to the channel conditions.

These contributions have been demonstrated in tele-echocardiography systems, leading to the following accomplished results:

- For the proposed compression format for storage purposes, the compression results depend on the acquisition devices, how the image is displayed, and the compression quality. However, the compression ratios obtained for the proposed format ranged from 19 to 41 and a compression gain of up to 75 % compared to JPEG 2000 was achieved without losing diagnostic information.

- For the proposed compression for real-time transmission purposes, compression and transmission rate recommendations have been given for each region. Minimum transmission rates have been recommended for each operation mode and the proposed echocardiogram compression technique in order to guarantee suitable clinical quality for transmission and storage using the proposed evaluation methodology. The recommended transmission rates for the ultrasound regions are the following: 200 Kbps for the 2-D and the color Doppler modes, and 40 Kbps for the M and the pulsed/continuous Doppler modes. These results make possible the transmission of echocardiogram videos over 3G wireless networks and beyond. The recommended transmission rates for the previous ultrasound transmission systems are considerably higher than those given for the proposed method which allows better transmission results. The proposed method requires less than 26% of the amount of data for the B mode, 78% for the Color Doppler mode, 5.2% for the M mode, and 16% for the pulsed/continuous Doppler mode compared with previous works.

- The use of ETP allows a saving in the transmitted bandwidth since the visualization characteristics are taken into account in the compression and transmission of the echocardiograms. The saving of bandwidth depends on the mode distribution of the echocardiograms. The percentage of time with adequate clinical quality increases between 5 % and 19 % for a mobile network with WiMAX access, representative settings, nine available echocardiograms and without an error control method.

- The selected configuration for SECM has been demonstrated to adapt to the channel conditions. Furthermore, the echocardiograms are visualized with adequate clinical quality for a mobile network with WiMAX access, representative settings and nine available echocardiograms. However, without SECM, the available echocardiograms are visualized without guaranteed clinical quality between 16.39 % and 25.46 % of the time for the same network.

- The proposed overall system for real-time echocardiogram transmission over wireless networks achieves a saving in the transmitted bandwidth while ensuring all the available echocardiograms are received with adequate clinical quality. This bandwidth saving depends on the mode distributions of the echocardiogram and on the error distribution. For a mobile network with WiMAX access and representative settings, and the available echocardiograms, the transmitted bandwidth is between 154 kbps and 244 kbps. If the results are compared with previous works where clinical videos with similar resolution are transmitted over WiMAX networks, the saving in transmitted bandwidth is higher than 1 Mbps. This bandwidth saving leads to savings in energy, money and transmission time.

5.3 Future Work

Although the objectives have been fulfilled and very good results have been obtained, slight improvements can be made in order to simplify the set up process. Moreover, the proposed methods can be used in other fields and can also be extended to other previously developed clinical image techniques as well as new techniques arising in the future.

- The configuration parameters of ETP and SECM have been set for a good performance using a WiMAX network with variant channel conditions. Although this configuration is expected to fit other networks, other parameters may be better for them. An automatic adjustment of the buffer time will be useful. The optimization of other parameters is more dependent on the type of access network, while the buffer time optimization is more dependent on the whole network configuration and the particular conditions for each transmission. This adjustment can be performed by sending some packets to test the network delay and the recovery time prior to transmission, after having selected the rest of the parameters.

- Simulations in other access networks can be carried out in order to set the appropriate parameters for each type of access network and to check if the obtained results are equivalent to the results obtained for the simulated network with WiMAX.

- A new scenario in which a multicast communication is established can be added. In this way, the captured echocardiogram could be transmitted to several destinations at the same time. In order to be able to make a multicast communication, a new application protocol must be designed. It is important to take into account that ETP uses communication over TCP and UDP to transmit the echocardiogram.

- The proposed echocardiogram storage format and the proposed compression method for echocardiogram transmission can be extended to other clinical image modalities in order to save storage space and to compress the images efficiently. The compression gain for other modalities depends on the image distribution of the different regions, but a saving in the storage space and transmission rate with respect to current compression formats and codecs is expected, as well as greater simplicity in the compression process.

- The evaluation methodology may be extended to other recent developments, such us 3-D echocardiography, which can be incorporated into the overall system. Moreover, the evaluation methodology can be used for any medical technique in order to use resources more efficiently and to use the minimal transmission rate or compression rate.

- The overall system for real-time echocardiogram transmission over wireless networks proposed in this Thesis can be applied to other clinical image modalities. The same steps and methods that have been followed in the design for echocardiograms may be adapted for other clinical videos.

Appendix A

Document Type Definition example for ETP

```
<?xml version="1.0" encoding="UTF-8"?>
<!ELEMENT configuration (tsize?,device?,delay?,end?,region+)>

<!ELEMENT tsize EMPTY>
<!ATTLIST tsize w CDATA #REQUIRED>
<!ATTLIST tsize h CDATA #REQUIRED>

<!ELEMENT device (#PCDATA)>

<!ELEMENT delay (#PCDATA)>

<!ELEMENT end EMPTY>

<!ELEMENT region ((ultr|vid|img|text|sound|ecg), position?, size?, codification?, mps?)>
<!ATTLIST region reliability CDATA #IMPLIED>
<!ATTLIST region id ID #REQUIRED>
<!ATTLIST region mode CDATA #REQUIRED>
<!ATTLIST region delete (yes|no) "no">

<!ELEMENT ultr (calibration?)>
<!ATTLIST ultr color (yes|no) "no">
<!ATTLIST ultr type (sweep|2D) "2D">
<!ATTLIST ultr fps CDATA #IMPLIED>
<!ATTLIST ultr pps CDATA #IMPLIED>
<!ATTLIST ultr bitrate CDATA #IMPLIED>
<!ATTLIST ultr video (stop|forward|back) #IMPLIED>
<!ATTLIST ultr resolution CDATA #IMPLIED>

<!ELEMENT calibration (t?,d?,v?)>
<!ELEMENT t (#PCDATA)>
<!ATTLIST t measurements CDATA #IMPLIED>
<!ELEMENT d (#PCDATA)>
<!ATTLIST d measurements CDATA #IMPLIED>
<!ELEMENT v (#PCDATA)>
<!ATTLIST v measurements CDATA #IMPLIED>
```

```
<!ELEMENT vid EMPTY>
<!ATTLIST vid color (yes|no) "no">
<!ATTLIST vid fps CDATA #IMPLIED>
<!ATTLIST vid bitrate CDATA #IMPLIED>
<!ATTLIST vid video (stop|forward|back) #IMPLIED>
<!ATTLIST vid resolution CDATA #IMPLIED>

<!ELEMENT img EMPTY>
<!ATTLIST img color (yes|no) "no">
<!ATTLIST img bpp CDATA #IMPLIED>
<!ATTLIST img codim CDATA #IMPLIED>

<!ELEMENT text (#PCDATA)>
<!ATTLIST text fsize (8|10|12) "8">

<!ELEMENT sound EMPTY>
<!ATTLIST sound configuration CDATA #IMPLIED>
<!ATTLIST sound channels CDATA #IMPLIED>
<!ATTLIST sound bitrate  CDATA #IMPLIED>
<!ATTLIST sound sampler CDATA #IMPLIED>

<!ELEMENT ecg EMPTY>
<!ATTLIST ecg cod (1D|2D) "2D">
<!ATTLIST ecg der CDATA #IMPLIED>
<!ATTLIST ecg block (128|256|512|1024) "128">
<!ATTLIST ecg fact (3|4|5|6) "4">
<!ATTLIST ecg sf CDATA #IMPLIED>
<!ATTLIST ecg time CDATA #IMPLIED>
<!ATTLIST ecg bpp CDATA #IMPLIED>
<!ATTLIST ecg ppslice CDATA #IMPLIED>
<!ATTLIST ecg width CDATA #IMPLIED>

<!ELEMENT position EMPTY>
<!ATTLIST position x CDATA #REQUIRED>
<!ATTLIST position y CDATA #REQUIRED>

<!ELEMENT size EMPTY>
<!ATTLIST size w CDATA #REQUIRED>
<!ATTLIST size h CDATA #REQUIRED>

<!ELEMENT codification (#PCDATA)>

<!ELEMENT mps (#PCDATA)>
```

Bibliography

[1] (2013) World Health Organization. [Online]. Available: http://www.who.int

[2] A. S. Go and et al., "Executive summary: heart disease and stroke statistics–2013 update: a report from the American Heart Association." vol. 127, no. 1, pp. 143–152, Jan. 2013.

[3] "Defunciones según la Causa de Muerte," 2013, instituto Nacional de Estadística (INE), http://www.ine.es/prensa/np767.pdf.

[4] H. Oh, C. Rizo, M. Enkin, and A. Jadad, "What is eHealth (3): a systematic review of published definitions," *J Med Internet Res*, vol. 7, no. e1, 2005.

[5] R. Wootton and J. Craig, *Introduction to Telemedicine: Royal Society of Medicine Press*, 1999.

[6] C. Pagliari, D. Sloan, P. Gregor, F. Sullivan, D. Detmer, J. P. Kahan, W. Oortwijn, and S. MacGillivray, "What is eHealth (4): a scoping exercise to map the field," *J Med Internet Res*, vol. 7, no. e9, 2005.

[7] A. G. Ekeland, A. Bowes, and S. Flottorp, "Effectiveness of telemedicine: a systematic review of reviews," *International Journal of Medical Informatics*, vol. 79, no. 11, pp. 736–71, 2010.

[8] C. Pattichis, E. Kyriacou, S. Voskarides, M. Pattichis, R. Istepanian, and C. Schizas, "Wireless telemedicine systems: an overview," *Antennas and Propagation Magazine, IEEE*, vol. 44, no. 2, pp. 143 –153, apr 2002.

[9] D. Giansanti and S. Morelli, "Digital tele-echocardiography: a look inside," *Ann Ist Super Sanita*, vol. 45, no. 4, pp. 357–62, 2009.

[10] G. Hooper, P. Yellowlees, T. Marwick, P. Currie, and B. Bidstrup, "Telehealth and the diagnosis and management of cardiac disease." *J Telemed Telecare*, vol. 7, no. 5, pp. 249–56, 2001.

[11] R. O. Bonow, D. L. Mann, D. P. Zipes, and P. Libby, *Braunwald's Heart Disease: A Textbook of Cardiovascular Medicine.* Elsevier, 2011.

[12] A. Evangelista and et al., "European Association of Echocardiography recommendations for standardization of performance, digital storage and reporting of echocardiographic studies," *European Journal of Echocardiography*, vol. 9, pp. 438–448, 2008.

[13] M. Picard and et al., "American Society of Echocardiography recommendations for quality echocardiography laboratory operations," *Journal of the American Society of Echocardiography*, vol. 24, no. 1, pp. 1 –10, January 2011.

[14] S. M. Bierig, D. Ehler, M. L. Knoll, and A. D. Waggoner, "Minimum standards for the cardiac sonographer: A position paper," *Journal of the American Society of Echocardiography*, vol. 19, no. 5, pp. 471–474, May 2006.

[15] M. Ukrit, A.Umamageswari, and Dr.G.R.Suresh, "Article: A Survey on Lossless Compression for Medical Images," *International Journal of Computer Applications*, vol. 31, no. 8, pp. 47–50, October 2011, published by Foundation of Computer Science, New York, USA.

[16] T. H. Karson, R. C. Zepp, S. Chandra, A. Morehead, and J. D. Thomas, "Digital storage of echocardiograms offers superior image quality to analog storage, even with 20:1 digital compression: results of the Digital Echo Record Access Study." *J Am Soc Echocardiogr*, vol. 9, no. 6, pp. 769–78. [Online]. Available: http://www.biomedsearch.com/nih/Digital-storage-echocardiograms-offers-superior/8943436.html

[17] F. Y. Han, B. Soong, D. Watson, and J. Whitehall, "Realtime fetal ultrasound by telemedicine in Queensland. A successful venture?" *J. Telemed. Telecare*, vol. 7, no. 2, pp. 7–11, December 2001.

[18] S. K. Yoo and et al., "Performance of a web-based, realtime, tele-ultrasound consultation system over high-speed commercial telecommunication lines," *J. Telemed. Telecare*, vol. 10, no. 3, pp. 175–179, June 2004.

[19] J. D. Thomas and M. L. Main, "Digital echocardiographic laboratory: where do we stand?" *J Am Soc Echocardiogr*, vol. 11, no. 10, pp. 978–83, 1998. [Online]. Available: http://www.biomedsearch.com/nih/Digital-echocardiographic-laboratory-where-do/9804104.html

[20] J. D. Thomas, D. B. Adams, S. DeVries, D. Ehler, N. Greenberg, M. Garcia, L. Ginzton, J. Gorcsan, A. S. Katz, A. Keller, B. Khandheria, K. B. Powers, C. Roszel, D. S. Rubenson, and J. Soble, "Guidelines and Recommendations for Digital Echocardiography," *Journal of the American Society of Echocardiography*, vol. 18, pp. 287–97, March 2005.

[21] H. Feigenbaum, "Digital recording, display, and storage of echocardiograms," *J Am Soc Echocardiogr*, vol. 1, no. 5, pp. 378–83, 1988.

[22] S. Mobarek, Y. Gilliland, A. Bernal, J. Murgo, and J. Cheirif, "Is a Full Digital Echocardiography Laboratory Feasible for Routine Daily Use?" *Echocardiography*, vol. 13, no. 5, pp. 473–482, 1996.

[23] "The Health Insurance Portability and Accountability Act (P.L.104-191)," 1996, Enacted by the U.S. Congress.

[24] "The Personal Information Protection and Electronic Document Act," 2000, Enacted in Canada for protection of health information against commercial use.

[25] "Ley Orgánica de Protección de Datos de carácter personal (Organic Law for Protection of Personal Data)," 1999, Enacted in Spain.

[26] A. Panayides, M. S. Pattichis, C. S. Pattichis, and A. Pitsillides, "A Tutorial for Emerging Wireless Medical Video Transmission Systems," *IEEE Antennas & Propagation Magazine*, 2011.

[27] R. Istepanian, S. Laxminarayan, and C. Pattichis, *M-Health: Emerging Mobile Health Systems*, ser. Topics in Biomedical Engineering. International Book Series. Springer, 2010. [Online]. Available: http://books.google.es/books?id=lZbBcQAACAAJ

[28] N. Herscovici, C. Christodoulou, E. Kyriacou, M. Pattichis, C. Pattichis, A. Panayides, and A. Pitsillides, "m-Health e-Emergency Systems: Current Status and Future Directions - [Wireless corner]," *IEEE Antennas and Propagation Magazine*, vol. 49, no. 1, pp. 216–231, Feb. 2007. [Online]. Available: http://dx.doi.org/10.1109/map.2007.371030

[29] A. Bhagat and M. Atique, "Medical images: Formats, compression techniques and DICOM image retrieval a survey," in *Devices, Circuits and Systems (ICDCS), 2012 International Conference on*, march 2012, pp. 172 –176.

[30] (2013) DICOM (Digital Imaging and Communications in Medicine). [Online]. Available: http://medical.nema.org/

[31] H. Feigenbaum, "Digital echocardiography." *Am J Cardiol*, vol. 86, no. 4A, pp. 2G–3G, 2000.

[32] F. Miller, A. Vandome, and J. McBrewster, *Digital Imaging and Communications in Medicine: Medical Imaging, File Format, Communications Protocol, Internet Protocol Suite, National Electrical Manufacturers Association, Picture Archiving and Communication System*. Alphascript Publishing, 2009. [Online]. Available: http://books.google.es/books?id=ytzYQgAACAAJ

[33] J. D. Thomas, "The DICOM Image Formatting Standard: what it means for echocardiographers." *J Am Soc Echocardiogr*, vol. 8, no. 3, pp. 319–27. [Online]. Available: http://www.biomedsearch.com/nih/DICOM-Image-Formatting-Standard-what/7640025.html

[34] C. Doukas and I. Maglogiannis, "Region of Interest Coding Techniques for Medical Image Compression," *Engineering in Medicine and Biology Magazine, IEEE*, vol. 26, no. 5, pp. 29 –35, sept.-oct. 2007.

[35] P. Akhtar, M. Bhatti, T. Ali, and M. Muqeet, "Significance of ROI Coding using MAXSHIFT Scaling applied on MRI Images in Teleradiology-Telemedicine," *Journal of Biomedical Science and Engineering*, no. 1, pp. 110–115, 2008.

[36] C. Christopoulos, J. Askelof, and M. Larsson, "Efficient methods for encoding regions of interest in the upcoming JPEG2000 still image coding standard," *Signal Processing Letters, IEEE*, vol. 7, no. 9, pp. 247–249, 2000.

[37] M. Ansari and R. Anand, "Context based medical image compression with application to ultrasound images," in *India Conference, 2008. INDICON 2008. Annual IEEE*, vol. 1, dec. 2008, pp. 28 –33.

[38] K. V. Sridhar, "Implementation of prioritised ROI coding for medical image archiving using JPEG2000," in *Signals and Electronic Systems, 2008. ICSES '08. International Conference on*, 2008, pp. 239–242.

[39] K. Hamamoto and T. Nishimura, "Basic investigation on medical ultrasonic echo image compression by JPEG2000-availability of wavelet transform and ROI method," in *Engineering in Medicine and Biology Society, 2001. Proceedings of the 23rd Annual International Conference of the IEEE*, vol. 3, 2001, pp. 2449–2452 vol.3.

[40] D. Taubman and M. Marcellin, "JPEG2000: standard for interactive imaging," *Proceedings of the IEEE*, vol. 90, no. 8, pp. 1336 – 1357, aug 2002.

[41] A. Said and W. Pearlman, "A new, fast, and efficient image codec based on set partitioning in hierarchical trees," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 6, no. 3, pp. 243 –250, jun 1996.

[42] D. H. Foos, E. Muka, R. M. Slone, B. J. Erickson, M. J. Flynn, D. A. Clunie, L. Hildebrand, K. S. Kohm, and S. S. Young, "JPEG 2000 compression of medical imagery," pp. 85–96, 2000. [Online]. Available: +http://dx.doi.org/10.1117/12.386390

[43] A. Agarwal, A. Rowberg, and Y. Kim, "Fast JPEG 2000 decoder and its use in medical imaging," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 7, no. 3, pp. 184–190, 2003.

[44] T. H. OH and R. BESAR, "Medical image compression using JPEG-2000 and JPEG: A comparison study," *Journal of Mechanics in Medicine and Biology*, vol. 02, no. 03n04, pp. 313–328, 2002. [Online]. Available: http://www.worldscientific.com/doi/abs/10.1142/S021951940200054X

[45] (2011) Digital Imaging and Communications in Medicine (DICOM) Part 3: Information Object Definitions. [Online]. Available: http://medical.nema.org/Dicom/2011/11_03pu.pdf

[46] M. Weinberger, G. Seroussi, and G. Sapiro, "The LOCO-I lossless image compression algorithm: principles and standardization into JPEG-LS," *Image Processing, IEEE Transactions on*, vol. 9, no. 8, pp. 1309 –1324, aug 2000.

[47] N. Tsapatsoulis, C. Loizou, and C. Pattichis, "Region of Interest Video Coding for Low bitrate Transmission of Carotid Ultrasound Videos over 3G Wireless Networks," in *Engineering in Medicine and Biology Society, 2007. EMBS 2007. 29th Annual International Conference of the IEEE*, 2007, pp. 3717–3720.

[48] S. P. Rao, N. S. Jayant, M. E. Stachura, E. Astapova, and A. Pearson-Shaver, "Delivering diagnostic quality video over mobile wireless networks for telemedicine," *Int. J. Telemedicine Appl.*, vol. 2009, pp. 1:1–1:9, Jan. 2009. [Online]. Available: http://dx.doi.org/10.1155/2009/406753

[49] M. G. Martini and C. T. E. R. Hewage, "Flexible Macroblock Ordering for Context-Aware Ultrasound Video Transmission over Mobile WiMAX," *International Journal of Telemedicine and Applications*, vol. 2010, p. 14, May 2010.

[50] A. Panayides, M. Pattichis, C. Pattichis, C. Loizou, M. Pantziaris, and A. Pitsillides, "Atherosclerotic Plaque Ultrasound Video Encoding, Wireless Transmission, and Quality Assessment Using H.264," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 15, no. 3, pp. 387 –397, may 2011.

[51] C. Debono, B. Micallef, N. Philip, A. Alinejad, R. Istepanian, and N. Amso, “Cross Layer Design for Optimised Region of Interest of Ultrasound Video Data over Mobile Wimax,” *Information Technology in Biomedicine, IEEE Transactions on*, vol. PP, no. 99, p. 1, 2012.

[52] A. Panayides, Z. Antoniou, Y. Mylonas, M. Pattichis, A. Pitsillides, and C. Pattichis, “High-Resolution, Low-delay, and Error-resilient Medical Ultrasound Video Communication Using H.264/AVC Over Mobile WiMAX Networks,” *Biomedical and Health Informatics, IEEE Journal of*, vol. PP, no. 99, pp. 1–1, 2013.

[53] (2013) The Moving Picture Experts Group website. [Online]. Available: http://mpeg.chiariglione.org/

[54] (2012) Home of Xvid codec project. [Online]. Available: http://www.xvid.org/

[55] D. Marpe, T. Wiegand, and G. Sullivan, “The H.264/MPEG4 advanced video coding standard and its applications,” *Communications Magazine, IEEE*, vol. 44, no. 8, pp. 134 –143, aug. 2006.

[56] A. Alesanco and et al., “A clinical distortion index for compressed echocardiogram evaluation: recommendations for Xvid codec,” *Physiological Measurement*, vol. 30, no. 5, pp. 429–440, June 2009.

[57] P. Pedersen, B. Dickson, and J. Chakareski, “Telemedicine applications of mobile ultrasound,” in *Multimedia Signal Processing, 2009. MMSP ’09. IEEE International Workshop on*, oct. 2009, pp. 1 –6.

[58] R. S. H. Istepanian, N. Y. Philip, and M. G. Martini, “Medical QoS provision based on reinforcement learning in ultrasound streaming over 3.5G wireless systems,” *IEEE J.Sel. A. Commun.*, vol. 27, no. 4, pp. 566–574, May 2009. [Online]. Available: http://dx.doi.org/10.1109/JSAC.2009.090517

[59] G. Sullivan, J. Ohm, W.-J. Han, T. Wiegand, and T. Wiegand, “Overview of the High Efficiency Video Coding (HEVC) Standard,” *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 22, no. 12, pp. 1649–1668, 2012.

[60] A. Panayides, Z. Antoniou, M. Pattichis, C. Pattichis, and A. Constantinides, “High efficiency video coding for ultrasound video communication in m-health systems,” in *Engineering in Medicine and Biology Society (EMBC), 2012 Annual International Conference of the IEEE*, 2012, pp. 2170–2173.

[61] A. Panayides, Z. Antoniou, M. Pattichis, and C. Pattichis, “The use of H.264/AVC and the emerging high efficiency video coding (HEVC) standard for developing wireless ultrasound video telemedicine systems,” in *Signals, Systems and Computers (ASILOMAR), 2012 Conference Record of the Forty Sixth Asilomar Conference on*, 2012, pp. 337–341.

[62] B.-J. Kim, Z. Xiong, and W. Pearlman, “Low bit-rate scalable video coding with 3-D set partitioning in hierarchical trees (3-D SPIHT),” *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 10, no. 8, pp. 1374 –1387, dec 2000.

[63] L. Kaur, R. Chauhan, and S. Saxena, “Performance improvement of the SPIHT coder based on statistics of medical ultrasound images in the wavelet domain.” *J Med Eng Technol*, vol. 29, no. 6, pp. 297–301.

[64] H. Lalgudi, A. Bilgin, M. Marcellin, and M. Nadar, “Compression of fMRI and ultrasound images using 4D SPIHT,” in *Image Processing, 2005. ICIP 2005. IEEE International Conference on*, vol. 2, 2005, pp. II–746–9.

[65] B. Ramakrishnan and N. Sriraam, “Compression of dicom images based on wavelets and spiht for telemedicine applications.”

[66] S. Cho, D. Kim, and W. A. Pearlman, “Lossless Compression of Volumetric Medical Images with Improved Three-Dimensional SPIHT Algorithm.” *J. Digital Imaging*, vol. 17, no. 1, pp. 57–63, 2004. [Online]. Available: http://dblp.uni-trier.de/db/journals/jdi/jdi17.html#ChoKP04

[67] B. Ramakrishnan and N. Sriraam, “Compression of DICOM images based on wavelets and SPIHT for telemedicine applications.”

[68] V. K. Bairagi and A. Sapkal, “Automated region-based hybrid compression for digital imaging and communications in medicine magnetic resonance imaging images for telemedicine applications,” *Science, Measurement Technology, IET*, vol. 6, no. 4, pp. 247–253, 2012.

[69] A. Kassim and W. S. Lee, “Embedded color image coding using SPIHT with partially linked spatial orientation trees,” *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 13, no. 2, pp. 203–206, 2003.

[70] J. Shapiro, ““Embedded image coding using zerotrees of wavelet coefficients”,” *Signal Processing, IEEE Transactions on*, vol. 41, no. 12, pp. 3445 –3462, dec 1993.

[71] Y. Chen and W. A. Pearlman, “Three-Dimensional Subband Coding of Video Using the Zero-Tree Method,” in *Proc. SPIE*, 1996, pp. 1302–1309.

[72] B.-J. Kim and W. Pearlman, “An embedded wavelet video coder using three-dimensional set partitioning in hierarchical trees (SPIHT),” in *Data Compression Conference, 1997. DCC ’97. Proceedings*, mar 1997, pp. 251 –260.

[73] R. K. Kouassi, J.-C. Devaux, P. Gouton, and M. Paindavoine, “Application of the karhunen-loeve transform for natural color images analysis,” in *Signals, Systems amp; Computers, 1997. Conference Record of the Thirty-First Asilomar Conference on*, vol. 2, 1997, pp. 1740–1744 vol.2.

[74] K. Shen and E. Delp, “Color image compression using an embedded rate scalable approach,” in *Image Processing, 1997. Proceedings., International Conference on*, vol. 3, 1997, pp. 34–37 vol.3.

[75] M. Saenz, P. Salama, K. Shen, E. Delp, M. Sáenz, P. Salama, K. Shen, and E. J. Delp, “An evaluation of color embedded wavelet image compression techniques,” in *SPIE Conference on Visual Communications and Image Processing 99*, 1999, pp. 282–293.

[76] R. S. D. W. B. M. Santhi, "Inter Color Correlation Based Enhanced Color SPIHT Coder," *European Journal of Scientific Research*, vol. 57, no. 4, pp. 592–600, 2011.

[77] M. Antonini, M. Barlaud, P. Mathieu, and I. Daubechies, "Image coding using wavelet transform," *IEEE Transactions on Image Processing*, vol. 1, no. 2, pp. 205–220, Apr. 1992. [Online]. Available: http://dx.doi.org/10.1109/83.136597

[78] J. Postel, "Transmission Control Protocol," RFC 793 (Standard), Internet Engineering Task Force, September 1981, updated by RFCs 1122, 3168. [Online]. Available: http://www.ietf.org/rfc/rfc793.txt

[79] ——, "User Datagram Protocol," Internet Engineering Task Force, RFC 768, August 1980. [Online]. Available: http://www.rfc-editor.org/rfc/rfc768.txt

[80] H. Schulzrinne, S. Casner, R. Frederick, and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications," RFC 3550, July 2008.

[81] G. Pelletier and K. Sandlund, "RObust Header Compression Version 2 (ROHCv2): Profiles for RTP, UDP, IP, ESP and UDP-Lite," RFC 5225 (Proposed Standard), Internet Engineering Task Force, April 2008. [Online]. Available: http://www.ietf.org/rfc/rfc5225.txt

[82] G. Pelletier, K. Sandlund, and L.-E. Jonsson, "RObust Header Compression (ROHC): A Profile for TCP/IP (ROHC-TCP," RFC 6846 (Proposed Standard), Internet Engineering Task Force, January 2013. [Online]. Available: http://tools.ietf.org/html/rfc6846

[83] T. Dierks and E. Rescorla, "The Transport Layer Security (TLS) protocol version 1.2." RFC 5246, 2008.

[84] Y. Wang, S. Wenger, J. Wen, and A. Katsaggelos, "Error resilient video coding techniques," *Signal Processing Magazine, IEEE*, vol. 17, no. 4, pp. 61–82, 2000.

[85] M. Schier and M. Welzl, "Optimizing Selective ARQ for H.264 Live Streaming: A Novel Method for Predicting Loss-Impact in Real Time," *Multimedia, IEEE Transactions on*, vol. 14, no. 2, pp. 415–430, 2012.

[86] ——, "Content-aware selective reliability for DCCP video streaming," in *Multimedia Computing and Information Technology (MCIT), 2010 International Conference on*, 2010, pp. 53–56.

[87] A. Huszák and S. Imre, "Source controlled semi-reliable multimedia streaming using selective retransmission in DCCP/IP networks," *Comput. Commun.*, vol. 31, no. 11, pp. 2676–2684, Jul. 2008. [Online]. Available: http://dx.doi.org/10.1016/j.comcom.2008.02.033

[88] L. Rizzo, "Effective erasure codes for reliable computer communication protocols," *SIGCOMM Comput. Commun. Rev.*, vol. 27, no. 2, pp. 24–36, Apr. 1997. [Online]. Available: http://doi.acm.org/10.1145/263876.263881

[89] X. Yang, C. Zhu, Z. G. Li, X. Lin, and N. Ling, "An unequal packet loss resilience scheme for video over the Internet," *Multimedia, IEEE Transactions on*, vol. 7, no. 4, pp. 753–765, 2005.

[90] N. Thomos, S. Argyropoulos, N. Boulgouris, and M. Strintzis, "Robust Transmission of H.264/AVC Video using Adaptive Slice Grouping and Unequal Error Protection," in *Multimedia and Expo, 2006 IEEE International Conference on*, 2006, pp. 593–596.

[91] S. Cho and W. A. Pearlman, "Multilayered protection of embedded video bitstreams over binary symmetric and packet erasure channels," *J. Vis. Comun. Image Represent.*, vol. 16, no. 3, pp. 359–378, Jun. 2005. [Online]. Available: http://dx.doi.org/10.1016/j.jvcir.2004.08.001

[92] P. Frossard, "FEC performance in multimedia streaming," *Communications Letters, IEEE*, vol. 5, no. 3, pp. 122–124, 2001.

[93] M. Chatterjee, S. Sengupta, and S. Ganguly, "Feedback-based real-time streaming over WiMax," *Wireless Communications, IEEE*, vol. 14, no. 1, pp. 64–71, 2007.

[94] A. Majumda, D. Sachs, I. Kozintsev, K. Ramchandran, and M. Yeung, "Multicast and unicast real-time video streaming over wireless LANs," *Circuits and Systems for Video Technology, IEEE Transactions on*, vol. 12, no. 6, pp. 524–534, 2002.

[95] A. Alinejad, N. Philip, and R. Istepanian, "Cross-Layer Ultrasound Video Streaming Over Mobile WiMAX and HSUPA Networks," *Information Technology in Biomedicine, IEEE Transactions on*, vol. 16, no. 1, pp. 31 –39, jan. 2012.

[96] 3GPP, "Overview of the GPP Release 4," V1.1.2, 2010. [Online]. Available: http://www.3gpp.org/ftp/Information/WORK_PLAN/Description_Releases/

[97] ——, "Overview of the GPP Release 5," V0.1.1, 2010. [Online]. Available: http://www.3gpp.org/ftp/Information/WORK_PLAN/Description_Releases/

[98] ——, "Overview of the GPP Release 6," V0.1.1, 2010. [Online]. Available: http://www.3gpp.org/ftp/Information/WORK_PLAN/Description_Releases/

[99] ——, "Overview of the GPP Release 7," V0.9.16, 2012. [Online]. Available: http://www.3gpp.org/ftp/Information/WORK_PLAN/Description_Releases/

[100] ——, "Overview of the GPP Release 8," V0.2.10, 2013. [Online]. Available: http://www.3gpp.org/ftp/Information/WORK_PLAN/Description_Releases/

[101] ——, "Overview of the GPP Release 9," V0.2.9, 2013. [Online]. Available: http://www.3gpp.org/ftp/Information/WORK_PLAN/Description_Releases/

[102] ——, "Overview of the GPP Release 10," V0.1.8, 2013. [Online]. Available: http://www.3gpp.org/ftp/Information/WORK_PLAN/Description_Releases/

[103] "IEEE standard for local and metropolitan area networks-Part 16: air interface for fixed broadband wireless access systems," *IEEE Std. 802.16e-2005*, May 2006.

[104] "IEEE standard for local and metropolitan area networks-Part 16: air interface for fixed broadband wireless access systems. Amendment 3: advanced air interface," *IEEE Std. 802.16m-2011*, May 2011.

[105] *IEEE Std. 802.16-2004: IEEE Standard for Local and metropolitan area networks Part 16: Air Interface for Fixed Broadband Wireless Access Systems*, Jun 2004.

[106] A. Ghosh, D. Wolter, J. Andrews, and R. Chen, "Broadband wireless access with WiMax/802.16: current performance benchmarks and future potential," *Communications Magazine, IEEE*, vol. 43, no. 2, pp. 129–136, 2005.

[107] K. Etemad, "Overview of mobile WiMAX technology and evolution," *Communications Magazine, IEEE*, vol. 46, no. 10, pp. 31–40, 2008.

[108] I. Papapanagiotou, D. Toumpakaris, J. Lee, and M. Devetsikiotis, "A survey on next generation mobile WiMAX networks: objectives, features and technical challenges," *Commun. Surveys Tuts.*, vol. 11, no. 4, pp. 3–18, Oct. 2009. [Online]. Available: http://dx.doi.org/10.1109/SURV.2009.090402

[109] R. K. Jha, A. W. Vishal, and U. D. Dalal, "A Survey of Mobile WiMAX IEEE 802.16m Standard," *CoRR*, vol. abs/1005.0976, 2010.

[110] D. Pareit, B. Lannoo, I. Moerman, and P. Demeester, "The History of WiMAX: A Complete Survey of the Evolution in Certification and Standardization for IEEE 802.16 and WiMAX," *Communications Surveys Tutorials, IEEE*, vol. 14, no. 4, pp. 1183–1211, 2012.

[111] A. Ghosh, D. R. Wolter, J. G. Andrews, and R. Chen, "Broadband wireless access with wimax/802.16: current performance benchmarks and future potential," *Comm. Mag.*, vol. 43, no. 2, pp. 129–136, Feb. 2005. [Online]. Available: http://dx.doi.org/10.1109/MCOM.2005.1391513

[112] S. Chikkerur, V. Sundaram, M. Reisslein, and L. Karam, "Objective Video Quality Assessment Methods: A Classification, Review, and Performance Comparison," *Broadcasting, IEEE Transactions on*, vol. 57, no. 2, pp. 165–182, 2011.

[113] Z. Wang, A. C. Bovik, H. R. Sheikh, S. Member, and E. P. Simoncelli, "Image quality assessment: From error measurement to structural similarity," *IEEE Trans. Image Processing*, vol. 13, pp. 600–612, 2004.

[114] Z. Wang, , Z. Wang, and A. C. Bovik, "Why Is Image Quality Assessment So Difficult?" in *in Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing*, 2002, pp. 3313–3316.

[115] Z. Wang, H. R. Sheikh, and A. C. Bovik, "Objective video quality assessment," in *IN THE HANDBOOK OF VIDEO DATABASES: DESIGN AND APPLICATIONS*. CRC Press, 2003, pp. 1041–1078.

[116] S. Garawi, R. S. Istepanian, and M. A. Abu-Rgheff, "3G wireless communications for mobile robotic tele-ultrasonography systems," *Comm. Mag.*, vol. 44, no. 4, pp. 91–96, Sep. 2006. [Online]. Available: http://dx.doi.org/10.1109/MCOM.2006.1632654

[117] A. Panayides, Z. Antoniou, M. Pattichis, and C. Pattichis, "The use of H.264/AVC and the emerging high efficiency video coding (HEVC) standard for developing wireless ultrasound video telemedicine systems," in *Signals, Systems and Computers (ASILOMAR), 2012 Conference Record of the Forty Sixth Asilomar Conference on*, 2012, pp. 337–341.

[118] E. Cavero, A. Alesanco, and J. Garcia, “Nueva propuesta de compresión para el almacenamiento de pruebas ecocardiográficas,” in *CASEIB 2011 - XXIX Congreso Anual de la Sociedad Española de Ingeniería Biomédica*, Nov. 2011.

[119] M. Ansari and R. Anand, “Context based medical image compression for ultrasound images with contextual set partitioning in hierarchical trees algorithm,” *Advances in Engineering Software*, vol. 40, no. 7, pp. 487 – 496, 2009. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0965997808001567

[120] T. Bray, J. Paoli, C. M. Sperberg-McQueen, E. Maler, and F. Yergeau, “Extensible Markup Language (XML) 1.0 (Fifth Edition),” November 2008, w3C Recommendation. [Online]. Available: http://www.w3.org/TR/xml/

[121] D. Bandon, P. Trolliard, A. Garcia, C. Lovis, A. Geissbühler, and J.-P. Vallée, “Building an enterprise-wide PACS for all diagnostic images,” *International Congress Series*, vol. 1268, no. 0, pp. 279 – 284, 2004, CARS 2004 - Computer Assisted Radiology and Surgery. Proceedings of the 18th International Congress and Exhibition. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0531513104005631

[122] W. X. W. Group, *The XML 1.0 Standard (5th Edition)*, W3C W3C Recommendation, 2008. [Online]. Available: http://www.network-theory.co.uk/w3c/xml/

[123] (2013) GOCR open-source character recognition. [Online]. Available: http://jocr.sourceforge.net/

[124] E. Cavero, A. Alesanco, and J. Garci ́a, “A new approach for echocardiogram compression based on display modes,” in *Information Technology and Applications in Biomedicine (ITAB), 2010 10th IEEE International Conference on*, 2010, pp. 1–4.

[125] A. Alesanco and J. Garcia, “Automatic Real-Time ECG Coding Methodology Guaranteeing Signal Interpretation Quality,” *Biomedical Engineering, IEEE Transactions on*, vol. 55, no. 11, pp. 2519 –2527, nov. 2008.

[126] J. Valin, V. K., and T. T., “Definition of the Opus Audio Codec,” RFC 6716, September 2012.

[127] F. Yergeau, “RFC 3629, UTF-8, a transformation format of ISO 10646,” http://tools.ietf.org/html/rfc3629, November 2003. [Online]. Available: http://tools.ietf.org/html/rfc3629

[128] E. Cavero, A. Alesanco, and J. Garcia, “Enhanced Protocol for Real-Time Transmission of Echocardiograms Over Wireless Channels,” *Biomedical Engineering, IEEE Transactions on*, vol. 59, no. 11, pp. 3212–3220, 2012.

[129] E. Cavero, A. Alesanco, L. Trajkovic, C. Pattichis, and J. Garcia, “Proposal of real-time echocardiogram transmission based on visualization modes with wimax access,” in *Computing in Cardiology (CinC), 2013*, Sept.

[130] S. B. Wicker, *Reed-Solomon Codes and Their Applications*. Piscataway, NJ, USA: IEEE Press, 1994.

[131] K. Fall and S. Floyd, "Simulation-based comparisons of tahoe, reno and sack tcp," *SIGCOMM Comput. Commun. Rev.*, vol. 26, no. 3, pp. 5–21, Jul. 1996. [Online]. Available: http://doi.acm.org/10.1145/235160.235162